NAC-0175_R0

Analysis of UCL Simulations at the Lognormal Distribution

Performance of the Chebyshev UCL Estimators and Improved Recommendation Rules

6 January 2022

Prepared by

NEPTUNE AND COMPANY, INC.
1435 Garrison St, Suite 201, Lakewood, CO 80215

1. Title: Analysis of UCL Simulations at the Lognormal Distribution		
2. Filename: Analysis UCL lognormal sims report_Neptune formatted.docx		
3. Description: The report analyzes of the results of simulations of upper confidence limits (UCL) for the mean of the lognormal distribution, using various UCL procedures, after generated lognormal data sets were filtered through the goodness-of-fit (GoF) rules recommended for ProUCL 5.2. The purpose was to get an approximate characterization of the behavior of various UCLs calculated by ProUCL in data that have been identified by the GoF rules as putatively lognormal. The analysis culminates with the identification of a simple set of rules for recommending the choice of UCLs in data that the GoF rules have identified as tentatively lognormal.		
	Name	Date
4. Originator	John H. Carson Jr.	6 Jan 2022
5. Reviewer	Paul Black	4 Jan 2022
6. Remarks		

CONTENTS

FIGURES

TABLES

ACRONYMS AND ABBREVIATIONS

CI	confidence interval
CoC	chemical of concern
CV	coefficient of variation
DQO	data quality objectives
GoF	goodness-of-fit
LOESS	locally estimated scatterplot smoothing
log SD	log-scale standard deviation
LOR	log of the odds ratio
MLE	maximum likelihood estimator
RelRMSE	relative Root Mean Squared Error
RPart	Recursive Partitioning
UCL	upper confidence limit
UCL_{MAL}	UCL type with minimum average loss
UCL_{MML}	UCL type with minimax loss

1.0 Introduction

This report presents a preliminary analysis of the results of simulations of upper confidence limits (UCL) for the mean of the lognormal distribution after data sets are filtered through the goodness-of-fit (GoF) rules recommended for ProUCL 5.2. The purpose of this is to get an approximate characterization of the behavior of various UCLs calculated by ProUCL in data that have been identified by the GoF rules as putatively lognormal. The analysis culminates with the identification of a simple set of rules for recommending the choice of UCLs in data that the GoF rules have identified as tentatively lognormal.

A UCL for the mean is the upper end of a one-sided confidence interval (CI) for the mean. The coverage of a CI is the frequency of the CI covering the true mean, synonymous with the confidence level. For a UCL, coverage is the frequency of over-estimating the true mean. Although a UCL is defined in terms of a one-sided CI, in environmental practice it is very often used as a conservative estimator of the mean concentration of a chemical of concern (CoC), which is intended to represent the consequence of "reasonable maximum exposure" to the CoC. The word "reasonable" must be emphasized here. The fact that risks based on UCLs are then summed in a risk assessment can be exceedingly conservative, but that issue will not be addressed here. UCLs for human health risk assessment are usually calculated for a confidence level of 95% (significance level of 5%). Estimators with low bias and low variance are desirable. In the risk assessment setting, value should be placed on:

- low expected frequency of underestimating the mean (roughly equal to 100% minus the confidence level),
- small positive bias,
- small average estimation errors, and
- small probability of large overestimation.

UCLs for the mean are also used to assess compliance with environmental standards. Coverage is perhaps more important in this setting. However, the consequences of exceeding a regulatory standard by a large amount are generally deemed to be severe. This again emphasizes the importance of ensuring that UCL estimators do not produce large overestimates or that they do so very infrequently.

2.0 Simulation

The simulation used in this study generated 10,000 replicate data sets for each lognormal distribution used. These distributions have a common mean of 100 and a wide range of coefficients of variation (CVs) (25 values from 0.1 to 20) covering behavior from very slightly skewed to highly skewed. Since these are all lognormal distributions, the CV determines the standard deviation of logs of the values and vice versa. The CV is used as an index parameter for the populations simulated in order to easily fit simulations for other distributions (and mixtures) into the same framework.

The sample sizes of the simulated data sets range from 5 to 1,000 with 47 different values. For each replicate, a sample of size 1,000 was generated as a parent sample from which samples of the various sizes required were selected for computation of summary statistics and UCLs. The

UCLs simulated include the Chebyshev 95% UCL, the Chebyshev 90% UCL, the H-UCL, the t-UCL, the skewed t-UCL, the adjusted Gamma UCL, Hall's bootstrap UCL, the bootstrap-t UCL and the BCa bootstrap UCL. The UCLs had a target coverage level of 95%, except for the Chebyshev 90% UCL. The results, which took several days to compute using parallel computing with up to 50 CPU cores, give an accurate characterization of the behavior of the UCLs calculated.

The computations were performed in R 4.1.1. The H-UCL was computed using the EnvStats library version 2.4.0.

The GoF selection rules, revised for ProUCL 5.2, are that data sets are treated as lognormal if:

- they are rejected as normal at level 0.01 by both the Shapiro-Wilks and Anderson-Darling goodness-of-fit tests,
- they are rejected as being from a Gamma distribution at level 0.05 by both the Anderson-Darling and Kolmogorov-Smirnov goodness-of-fit tests, and
- they are not rejected as lognormal at level 0.1 by either the Shapiro-Wilks or the Anderson-Darling goodness-of-fit test.

Data filtered according to these criteria were used to develop the performance measures and recommendation rules for UCLs for designated lognormal data in ProUCL 5.2. The ultimate objective is to model and improve the behavior of ProUCL 5.2, and later versions, when it is given approximately lognormal data.

3.0 UCL Properties

A CI (including one-sided CIs) for a parameter is defined as a random interval that covers the true (unknown) value of the parameter with specified probability (coverage probability or confidence coefficient) from the point of view of the procedure generating the random interval being applied over and over again to data generated by the same random process. In statistical literature, the two most important considerations for CI are coverage and accuracy. There is some disagreement about coverage. Most authors (such as Mood, Graybill, Boes, Randles, Wolfe, Hollander, C.R. Rao, Davison, and Hinkley) indicate the coverage probability should be at least approximately equal to the nominal value (confidence level) for a valid CI. A minority of authors (such as Bickel and Doksum) hold that any CI for which the coverage probability equals or exceeds the confidence level is valid. All of the authors listed have well-known books, listed in this report's References section, which discuss confidence intervals.

All of the authors listed above hold that accuracy is a very important consideration. Accuracy is a measure of the average closeness of the confidence limits to the true value of the parameter. In the case of two-sided confidence, the average length of the CI is universally held to be a good measure of accuracy. Generally, CIs are associated with the acceptance region of a statistical hypothesis test constructed for the parameter in question. Most powerful tests are associated with shortest length two-sided CIs and with most accurate one-sided confidence intervals. A UCL, U_1, is said to be more accurate than another UCL, U_2, if both UCLs have the same coverage probability, and on average $U_1 < U_2$. It is often easier to compare the averages of $U_1 \,/\, \mu$ to $U_2 \,/\, \mu$, which makes the results across different simulations or theoretical calculations easier to

compare. This suggests that estimation errors for UCLs be considered in terms of relative error, $U/\mu - 1$.

As far as Neptune and Company, Inc. (Neptune) is aware, accuracy of CIs has not previously (in version 5.1 and earlier) been considered in selecting methodologies for UCL computation in ProUCL nor in their recommendation in specific cases. Here are three examples that illustrate the importance of considering accuracy as well as coverage:

1. Suppose that one is attempting to estimate the average height of Americans, and measures the height of a few individual Americans. One could construct a UCL that is the observed sample average plus one inch. Or one could construct a UCL that is the observed sample average plus 10 feet. The first UCL is likely to be somewhat close to the true average, but its coverage probability is unknown. The latter UCL will have exceedingly good coverage probability (100%, as there are no individuals over the height of 10 feet). However, the latter UCL is clearly a ridiculous overestimate. It may be clear that it is an overestimate in this example, but were soil concentrations the measured variable instead, it might not be so obvious that this was a ridiculous overestimate.
2. Consider a UCL for mean concentration that is constructed as follows: all data is ignored. One instead uses a random number generator so that, 95% of the time, the UCL is chosen to be one million parts per million, and, the other 5% of the time, the UCL is chosen to be zero parts per million. This UCL guarantees *exactly* 95% coverage probability. However, when it overestimates the mean, it is likely to be a gross overestimation, and when it underestimates the mean, it is likely to be a gross underestimation. Few would consider this UCL to be a good method, despite its ideal coverage probability.
3. Suppose there were a UCL method that could construct an estimate that was equal to *exactly* 0.9999999 times the true mean every single time. Such a UCL method would have 0% coverage probability, and yet would lead to an extremely accurate estimate of the true mean.

These examples suggest that coverage probability alone is a poor metric by which to judge a UCL; accuracy should be considered as well. A slight underestimate is not necessarily worse than a gross overestimate; some balance is needed. This would better align ProUCL with the data quality objectives (DQO) process, which addresses both false positive and negative error rates. When dealing with human health and ecological risks, there is a desire to be conservative, but there are limits to this, as extreme conservatism could lead to expending great resources at a site that does not require it, and preventing those resources from being utilized to better effect elsewhere. Further, since risk calculations are generally calculated as linear effects with respect to the mean concentration (represented by the UCL), slight overestimates or underestimates in mean concentrations do not have a compounding effect on subsequent risk calculations. Furthermore, the consequences of exceeding a regulatory standard are generally deemed to be increasingly severe with increasing magnitude of the exceedance.

The considerations above can be formulated into loss functions, and UCL procedures can be selected or ranked based on their risk (that is, their expected or long run average performance with respect to the relevant loss functions). The use of loss functions in the face of uncertainty to choose actions that minimize risk (expected loss) was developed in a branch of mathematics known as Decision Theory. Decision Theory is used extensively in many fields of practice, such

as business, economics[1], applied mathematics, statistics, computer science, and engineering. According to Winkler (1972), writing from the decision theoretic point of view, "[a]n optimal interval estimate … minimizes the decision maker's expected loss." Lehmann (1986) defines a uniformly most accurate lower confidence bound, and analogously a uniformly most accurate upper confidence bound (limit), as the estimator that minimizes a specified estimation loss function while satisfying the desire coverage.

Loss and risk functions can be fine-tuned to reflect the requirements of a very specific application of a UCL. Conversely, a range of loss and risk functions can be used to show that a UCL procedure is robust for a range of applications. This is the approach taken to evaluate UCL performance in ProUCL and is the motivation for the loss functions described in the following sections.

3.1 Bounded Loss for UCLs

Following Casella, Hwang, and Robert (1990), Casella and Hwang (1991), and Casella, Hwang, and Robert (1993), bounded linear loss functions for UCLs are the sum of a size function that penalizes inaccuracy and the 0-1 loss that penalizes lack of coverage. The average of the 0-1 loss over a sample equals 1 minus the empirical coverage, so it is a penalty for lack of coverage. In the case of one-sided intervals, specifically UCLs with 0 as a lower bound, the size of the interval measured as its length equals the UCL. Therefore, it makes sense to consider an alternative size measure that measures the closeness of the UCL to the true mean. An even function of the relative difference of the UCL from the true mean works well. The associated risk is the average of the loss function over a population or over a large sample, such as one generated by simulation.

One pair of bounded linear loss and risk functions, using absolute relative error in the rational size function defined in Casella, Hwang, and Robert (1990, page 10), is given by:

[1] In business and economics, one maximizes expected utility, with utility being the negative of loss.

$$
\begin{aligned}
\mu &= \text{true mean} \\
u_i &= \text{UCL}_i, \text{ for } i = 1, \ldots, N, \text{ are UCL estimates} \\
q_i &= I(u_i < \mu) = \begin{cases} 0, & u_i \geq \mu \\ 1 & u_i < \mu \end{cases} \\
v_i &= \left| \frac{u_i - \mu}{\mu} \right| = \left| \frac{u_i}{\mu} - 1 \right| \\
L_B(u_i|\mu, a) &= \frac{v_i}{a + v_i} + q_i \\
1 - \bar{q} &= 1 - \frac{1}{N} \sum_{i=1}^{N} q_i \text{ is the empirical coverage} \\
U &= (u_1, \ldots, u_N) \\
R_B(U|\mu, a) &= \frac{1}{N} \sum_{i=1}^{N} \frac{v_i}{a + v_i} + \bar{q},
\end{aligned}
$$

where $I(u < \mu)$ is the indicator function that equals *1* if $u < \mu$ and *0* otherwise, q_i is the 0-1 loss, v_i is the absolute relative error of the UCL estimate, $\bar{q}$ is the average of the 0-1 loss (an estimate of the probability that the interval does not cover the true value), $L_B(u_i|\mu, a)$ is the loss incurred by an individual UCL estimate, and $R_B(U|\mu, a)$ is the expected loss or the loss averaged over a population or a large sample. The use of relative absolute error in the loss makes it consistent and easy to compare for different values of the true mean.

This particular loss function has some interesting and useful theoretical properties when used for estimating two-sided intervals, as discussed in several papers, including Casella, Hwang, and Robert (1990), Casella and Hwang (1991), and Casella, Hwang, and Robert (1993). It is important that the penalty for lack of coverage and the penalty for inaccuracy be approximately balanced so that neither dominates the other. When these are not balanced, problems like Berger's paradox can occur (for example, Casella, Hwang, and Robert (1993)). In the bounded loss shown above, the loss is 0 when the UCL equals the true mean. The loss increases to $1 + \frac{1}{1+a}$ as the UCL decreases from the true mean to 0 and increases to the limit of 1 as UCL increases.

However, note that this loss does not penalize over-coverage. Given that many authors regard substantial over-coverage as a negative (see Section 3.0), this is a less than desirable feature. An alternative would be to have penalties for both under-coverage and over-coverage that are appropriate for a given application. This will be explored further in the section on unbounded loss (Section 3.2).

Since for typical applications we want a UCL to be a conservative estimator but not overly so, that is, to have both good coverage and good accuracy, it makes sense to penalize underestimates of the true mean more than overestimates. This gives weighted bounded linear loss and risk functions:

$$L_{wB}(U|\mu, a, b_-, b_+) = [b_- \cdot q_i + b_+ \cdot (1 - q_i)] \frac{v_i}{a + v_i} + q_i,$$

$$R_{wB}(U|\mu, a, b_-, b_+) = \frac{1}{N} \sum_{i=1}^{N} \left\{ [b_- \cdot q_i + b_+ \cdot (1 - q_i)] \frac{v_i}{a + v_i} \right\} + \bar{q},$$

where b_-, b_+ are low and high bias penalty coefficients. If $b_- = b_+ = 1$, then this is the same as the previous loss function. As the UCL estimate goes to 0 from μ, the loss increases from 0 to $1 + b_- \cdot \frac{1}{1+a}$, and as the UCL estimate increases from μ, the loss increases from 0 to b_+. Setting $b_+ = 1 + b_- \cdot \frac{1}{1+a}$ makes the losses equal in the cases of maximum underestimation and maximum overestimation. This is the balanced case for a mean that cannot be negative.

Figure 1 illustrates the weighted bounded linear loss function, L_{Bw}, given above with the true value $\mu = 100$, $b_- = 1$ and various values for b_+ and a. Plots labelled "balanced" have the parameters balanced in the sense given in the previous paragraph.

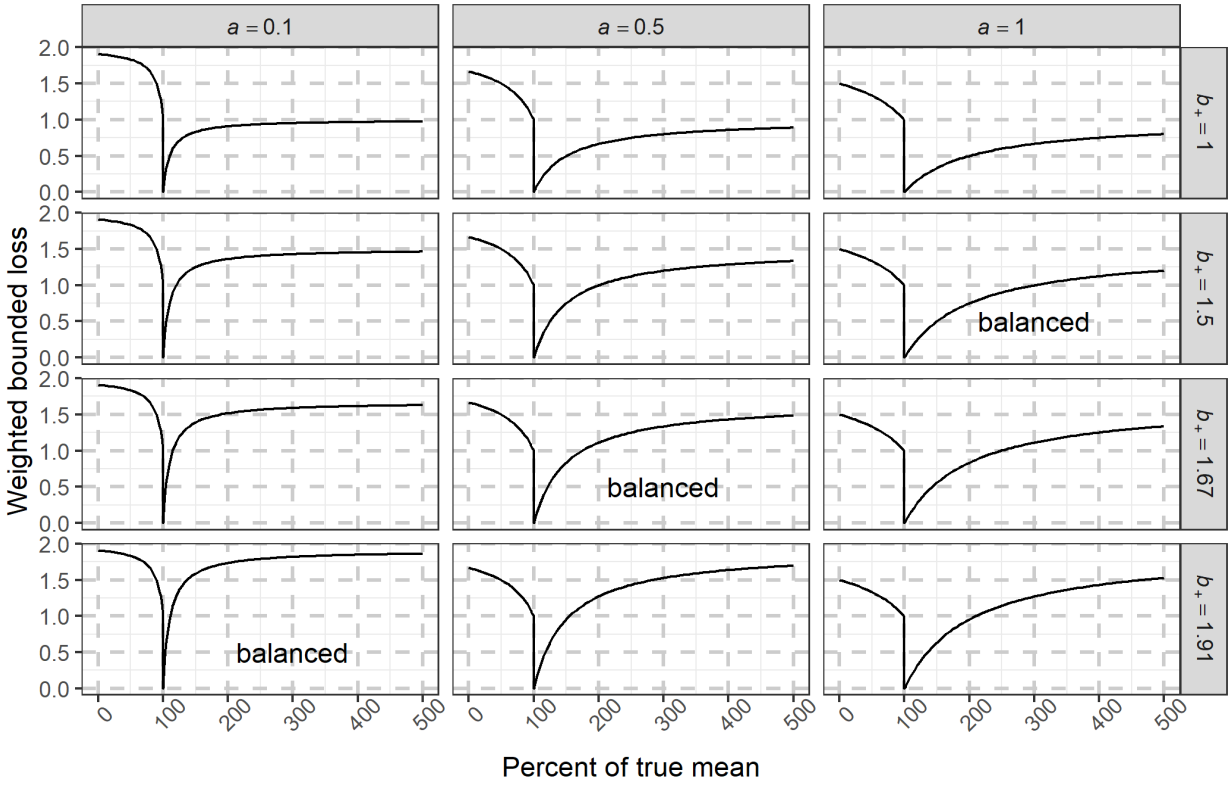

Figure 1. Weighted linear bounded loss profiles for various parameter values. Values of b_+ are chosen as: i) $b_+ = b_- = 1$, ii) $b_+ = 1.5$ matches the maximum loss for under- and overestimation when $a = 1$; iii) $b_+ = 1.67$ matches the maximum loss for under- and overestimation when $a = 0.5$, and iv) $b_+ = 1.91$ matches the maximum loss for under- and overestimation when $a = 0.1$.

Review of the plots in the first row of Figure 1 leads to some interesting questions. When it comes to assessment of risk or determining whether a site meets cleanup criteria, should a UCL underestimating a mean by 1% be penalized more than a UCL overestimating the true mean by 200% or 500% or 1,000%? This will always be the case if $b_+ \leq 1$, as in the first row of plots. The balanced value of $b_+ = 1 + b_- \cdot \frac{1}{1+a}$, seen in the plots labeled "balanced," ensures that this can never happen.

To further explain the examples from the plots above, when $a = 0.1, b_- = 1, b_+ = 1.91$, a 1% underestimate of the mean is penalized approximately the same as an 11% overestimate of the mean. When $a = 0.5, b_- = 1, b_+ = 1.67$, a 1% underestimate of the mean is penalized approximately the same as a 75% overestimate of the mean. When $a = 1, b_- = 1, b_+ = 1.5$, a 1% underestimate of the mean is penalized approximately the same as a 200% overestimate of the mean. Therefore, minimizing the bounded linear loss with $a = 1,\ b_- = 1, b_+ = 1.5$, gives a much more conservative (larger) UCL than does minimizing the bounded linear loss with $a = 0.1,\ b_- = 1, b_+ = 1.91$, which gives more accurate estimates. Using $a = 1,\ b_- = 1, b_+ = 1.5$ is a compromise between conservatism and accuracy. In Section 4.0, UCL Plots, these losses are designated as "Loss_bnd_c," "Loss_bnd_a," and "Loss_bnd_m," with "c," "a," and "m" indicating conservative, accurate, and intermediate, respectively.

3.2 Unbounded Loss for UCLs

There are many possibilities for unbounded loss functions for confidence intervals, including UCLs. One of the general advantages of unbounded loss functions over bounded loss functions is that bounded loss functions for estimators which are unbounded (at least for practical purposes) are forced to have the left and right sections of the loss function be concave, as may be seen in Figure 1 above. On the other hand, unbounded loss functions can be convex. The graphs of concave functions bend downward, and those of convex functions bend upward. Convex loss functions generally have important properties for estimation procedures, as discussed at length by Lehmann (1983) and specifically for interval estimates by Winkler (1972). Lehmann shows that, under strictly convex loss functions, there is always an essentially unique estimator that achieves minimum risk, and any estimator that does this is a function of the sufficient statistics[2] of the distribution of the data. This is the Rao-Blackwell theorem. This extremely important result does not hold for concave loss functions. The Gamma, t-, and H-UCL estimators in ProUCL are functions of the sufficient statistics for the Gamma, normal, and lognormal distributions, respectively. Many other results concerning convex loss functions in Lehmann (1983) are useful but highly technical and will not be discussed further here.

Although many properties of optimal point estimators do not carry over to interval estimators, Winkler (1972) shows that, when a strictly convex loss function is applied to the selection of an interval estimator given a specific distribution, there is an optimal interval estimator, which is not guaranteed for a concave loss function or a convex loss function that is not strictly convex. Absolute error loss is an example of convex loss function that is not strictly convex. While strict

[2] The sufficient statistics are functions of the data that for a specific distribution summarize all relevant information contained in the sample about the parameters of the distribution. As an example, the sample mean and variance are the sufficient statistics for the normal distribution.

convexity of the loss function does not guarantee optimality in the general situation of data from an unknown distribution, it indicates that using strictly convex loss functions to evaluate candidate UCL procedures is a productive approach.

The unbounded loss functions used here for evaluating UCLs are composed of four parts. The first two are penalties for under- and over-coverage. The second two are penalties for inaccuracy (the combination of bias and imprecision), on the low side and on the high side. As with bounded loss functions, it is important that the coverage and accuracy components of the loss be balanced so that both contribute meaningfully. This balance is first provided by choosing component losses that are 0 when the UCL equals the true mean and that increase without bound as the UCL is further and further from the true value and as the UCL coverage goes to 0 or to 1. Secondly, the rates of increase for the negative and positive elements of the component losses can be varied to reflect the losses incurred by the respective errors.

3.2.1 Unbounded Coverage Loss for UCLs

There are a couple of very reasonable possibilities for the unbounded coverage loss. One is based on the log of the odds ratio (LOR) of the expected coverage of the UCL estimator versus the desired or target coverage level (95% in our case). The LOR is a very natural scaling to compare probabilities (or coverages), and its absolute value is a natural distance metric between probabilities. We use the square of the LOR as a loss function:

$$\begin{aligned}
\gamma &= \text{desired coverage of UCL} \\
c_- > c_+ &\quad \text{are} \quad \text{low and high coverage penalty coefficients} \\
\text{logit}(p) &= \log\left(\frac{p}{1-p}\right) \\
\text{LOR}(p|\gamma) &= \log\left(\frac{p}{1-p}\cdot\frac{1-\gamma}{\gamma}\right) = \text{logit}(p) - \text{logit}(\gamma) \\
L_{\text{LOR}}(\bar{q}|\gamma) &= LOR(\bar{q}|\gamma)^2 \\
L_{\text{wLOR}}(\bar{q}|\gamma, c_-, c_+) &= [c_- \cdot I(\bar{q} \le \gamma) + c_+ \cdot I(\bar{q} > \gamma)] \cdot LOR(\bar{q}|\gamma)^2
\end{aligned}$$

Here L_{wLOR} is the weighted version of the LOR coverage loss function. Both are strictly convex loss functions.

Another natural metric depends on the fact that the distribution of UCL estimators is asymptotically normal. This is a result of a number of things coming into play: the use of maximum likelihood estimators (MLEs) in computing UCLs, the asymptotic normality of MLEs under regularity conditions, UCLs being continuous functions of MLEs, and various convergence results (see Serfling (1980)). This is not difficult to show. This fact then suggests that coverage probabilities of UCL estimators could usefully be compared to the desired coverages in the probit scale, the inverse of the normal probability function:

$$\begin{aligned}
L_{\Phi}(\bar{q}, |\gamma) &= 3.763 \cdot [\Phi^{-1}(\bar{q}) - \Phi^{-1}(\gamma)]^2 \\
L_{w\Phi}(\bar{q}|\gamma, c_-, c_+) &= 3.763 \cdot [c_- \cdot I(\bar{q} \le \gamma) + c_+ \cdot I(\bar{q} > \gamma)] \cdot [\Phi^{-1}(\bar{q}) - \Phi^{-1}(\gamma)]^2
\end{aligned}$$

where Φ^{-1} is the quantile function of the standard normal distribution, 3.763 is a scaling factor to match the LOR and probit losses for coverage of 0.8 versus the target level of 0.95. $L_{w\Phi}$ is the weighted form of the probit coverage loss. Both are strictly convex loss functions.

Note that, unlike the bounded (0-1) coverage loss, which can be computed from individual UCL estimates, the unbounded coverage loss functions use the average empirical coverage from simulation. If the coverage were computed from theoretical calculations or from a large simulation, we would consider the loss calculated from the coverage to be an expected loss or risk.

This coverage is converted to a score (logit or probit) for comparison to the desired coverage. The target coverage (say 95%) is also converted to a logit (probit). Then we square the difference between them. If the target coverage logit (probit) is less than the true coverage logit (probit), this is under-coverage, and we weight the squared difference with a weight of c_{-}. For over-coverage, the squared difference is weighted by c_{+}. Generally, we weight under-coverage errors more than over-coverage errors but do give positive weight to over-coverage errors. This is consistent with the views on CI coverage of most of the authors surveyed (see Section 3.0, UCL Properties).

Also, larger coverage errors (in logit/probit scale) should be weighted much more than relatively small coverage errors instead of the weight being proportional to the size of the error. This is technically known as having a strictly convex loss function. Using the square of the error makes this a strictly convex loss function. Weighting a convex loss according to whether errors are under- or over-coverage errors, provided the weights are positive, also results in a convex loss function.

The shapes of the squared LOR and squared difference of probits coverage loss functions are shown in Figure 2 and Figure 3. Note that, since the loss is unbounded on both sides of the target coverage, γ, the loss must increase very steeply if the x-axis is scaled in probability. In the plots below, in which the x-axis is labelled as probability but scaled as a normal deviate (just as in a normal Q-Q plot), the loss function would appear symmetric, except that we are weighting under-coverage much more than over-coverage.

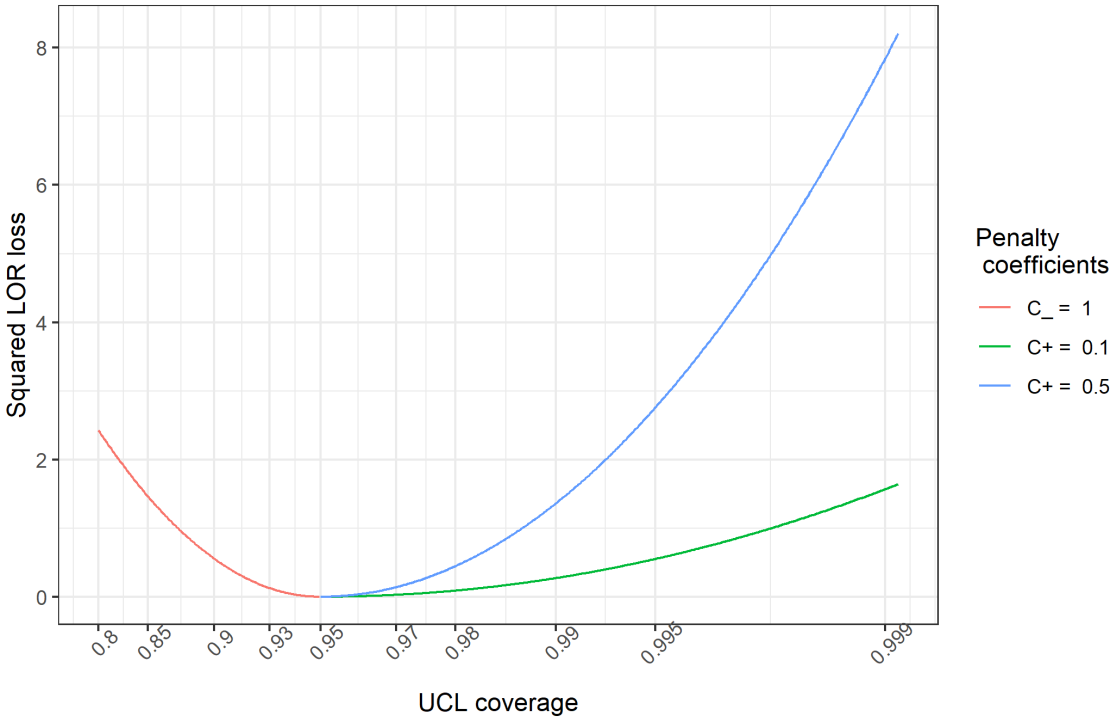

Figure 2. Weighted logit squared error loss examples for coverage. c_- = under-coverage penalty coefficient. c_+ = over-coverage penalty coefficient. Two different values of c_+ are shown.

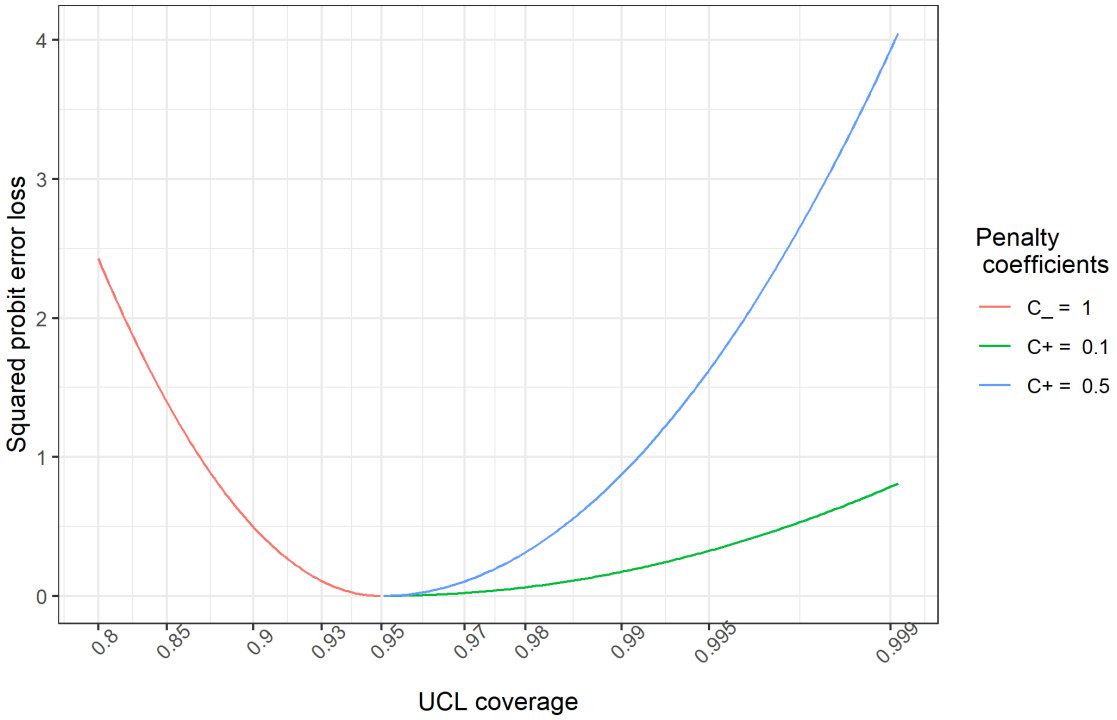

Figure 3. Weighted probit squared error loss examples for coverage. C_- = under-coverage penalty coefficient. C_+ = over-coverage penalty coefficient. Two different values of c_+ are shown.

The weighted squared LOR and weighted probit squared error loss functions are clearly similar but not identical.

3.2.2 Unbounded Accuracy Loss for UCLs

The penalty for lack of accuracy is relative mean squared error from the true mean value, which in these simulations is 100. This allows accounting for both the relative bias and the relative variance of the UCL as an estimator. For a UCL, negative deviations should be weighted more than positive deviations, since our objective is a conservative estimate of the mean. Larger error deviations should be penalized much more than smaller ones. Use of weighted squared error loss, as discussed above, minimizes the worst behavior of an estimator by penalizing large errors by much more than the magnitude of the error. This seems appropriate, because a major concern with UCLs in environmental applications has been the fact that some UCL procedures can produce wild overestimates of the mean under certain conditions. Having a very small penalty for small overestimates and a very large penalty for large overestimates results from the proposed weighted squared error loss and promotes our objective for a UCL that is a conservative, but not overly conservative, estimator of the mean.

$$
\begin{aligned}
q_i &= I(u_i < \mu) \\
v_i &= \left|\frac{u_i - \mu}{\mu}\right| = \left|\frac{u_i}{\mu} - 1\right| \\
L_{\text{wMSE}}(u_i|\mu, c_-, c_+) &= [b_- q_i + b_+(1 - q_i)]v_i^2 \\
R_{\text{wMSE}}(U|\mu, c_-, c_+) &= \frac{1}{N}\sum_{i=1}^{N}[b_- q_i + b_+(1 - q_i)]\, v_i^2
\end{aligned}
$$

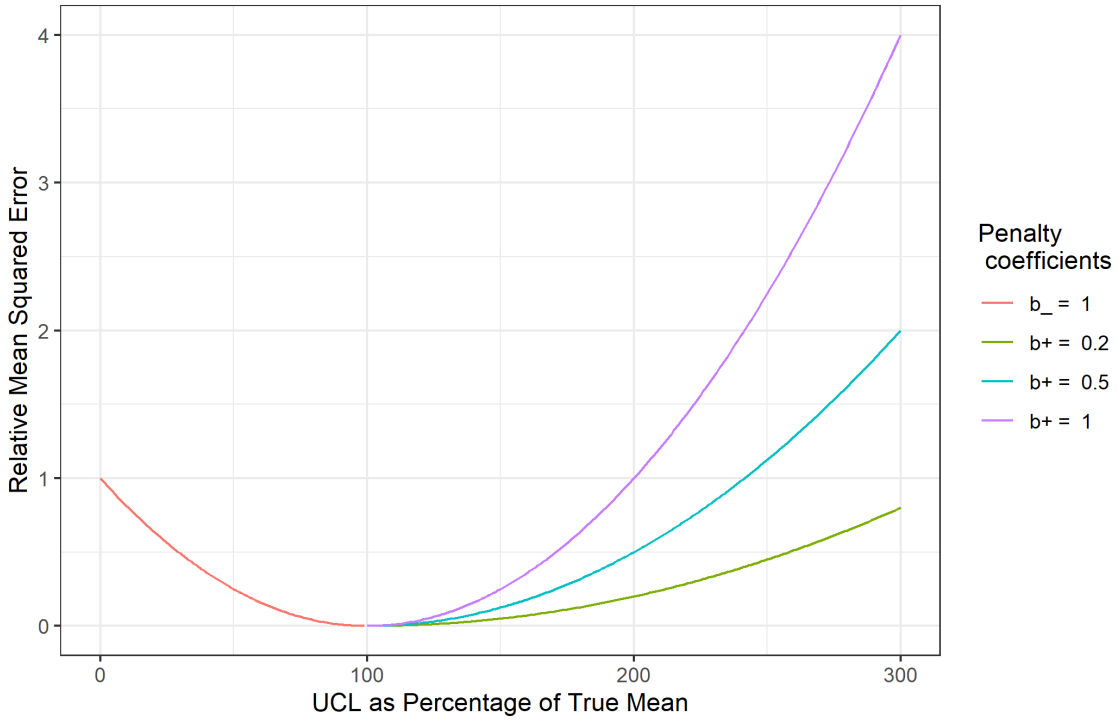

Figure 4. Weighted mean squared error loss examples for inaccuracy (the combination of bias and imprecision). B_ = negative bias penalty coefficient. B+ = positive bias penalty coefficient. Three different values of b+ are shown, corresponding to conservative, intermediate, and accurate estimates.

3.2.3 Combined Unbounded Loss

These components of the loss function, penalties for under- and over-coverage and for under- and over-estimation, are added together to create the weighted linear unbounded loss function.

The formulas are:

$$L_{wU,LOR}(U \mid \mu, \gamma, b_-, b_+, c_-, c_+) = L_{wMSE}(U \mid \mu, b_-, b_+) + L_{wLOR}(\bar{q} \mid \gamma, c_-, c_+)$$
$$L_{wU,\Phi}(U \mid \mu, \gamma, b_-, b_+, c_-, c_+) = L_{wMSE}(U \mid \mu, b_-, b_+) + L_{w\Phi}(\bar{q} \mid \gamma, c_-, c_+)$$

4.0 UCL Plots

The most important UCL estimators (omitting the Gamma approximate UCL, the jackknife UCL, and the percentile bootstrap UCL) computed in ProUCL are compared by looking at their performance with respect to several measures of performance, including coverage, relative bias (bias divided by the true value), variance of relative estimation error, relative Root Mean Squared Error (RelRMSE), and various bounded and unbounded loss functions designed to focus on different aspects of performance of the UCLs being compared.

The parameter values for each loss function in the plots below are summarized in Table 1. The names of the loss functions briefly indicate their characteristics. As discussed in Section 3.1

above, “Loss_bnd_c,” “Loss_bnd_m,” and “Loss_bnd_a” indicate weighted bounded loss with parameters chosen to be conservative, intermediate, and accurate, respectively.

Table 1. Loss function parameters

Name of Loss	Type of Loss	Type of Coverage Loss	a	b_-	b_+	c_-	c_+
Loss_bnd_c	Bounded	0-1	1	1	1.5	1	0.0
Loss_bnd_m	Bounded	0-1	0.5	1	1.67	1	0.0
Loss_bnd_a	Bounded	0-1	0.1	1	1.91	1	0.0
Loss_LOR_c_RelMSE_c	Unbounded	LOR	-	1	0.2	1	0.1
Loss_LOR_a_RelMSE_c	Unbounded	LOR	-	1	0.2	1	0.5
Loss_LOR_c_RelMSE_a	Unbounded	LOR	-	1	1.0	1	0.1
Loss_LOR_a_RelMSE_a	Unbounded	LOR	-	1	1.0	1	0.5
Loss_probit_c_RelMSE_c	Unbounded	Probit	-	1	0.2	1	0.1
Loss_probit_a_RelMSE_c	Unbounded	Probit	-	1	0.2	1	0.5
Loss_probit_c_RelMSE_a	Unbounded	Probit	-	1	1.0	1	0.1
Loss_probit_a_RelMSE_a	Unbounded	Probit	-	1	1.0	1	0.5

For the unbounded losses, “Loss_LOR_c_RelMSE_c,” “Loss_LOR_c_RelMSE_a,” “Loss_probit_c_RelMSE_c,” and “Loss_probit_c_RelMSE_a” have minimal over-coverage penalty because their coverage loss is “conservative.” The strings “LOR” and “probit” in the unbounded loss names refer to using either LOR or probit losses for coverage. The loss names ending in “c,” like “Loss_LOR_c_RelMSE_c,” indicate minimal loss for overestimation, resulting in a more conservative estimate. The loss names ending in “a,” like “Loss_LOR_c_RelMSE_a,” indicate symmetric loss for under- and overestimation, resulting in a more accurate estimate.

The comparisons are plotted graphically in Figure 5 through Figure 20 below, and patterns in the plots are explored and interpreted. The plots are organized by the log-scale standard deviation (log SD) of the individual simulated data sets. Although the data sets are generated based on specified values of the population CV (which is equivalent to specifying values of log SD, since they are functions of each other) and for various sample sizes, the computed log SDs vary by data set.

Furthermore, the GoF selection rules (Section 2.0) for filtering the generated data also somewhat change the distribution of log SD by sample size from what was originally generated. Each plot shows the features of the UCL estimators grouped by quartile of sample log SD (Figure 5

through Figure 20 in this section) and grouped by decile of the log SD (in Appendix A). The plots in Appendix A show more detail of the UCL behavior, since the plots are presented over a finer grid of sample log SD ranges.

Each combination of values of a type of UCL estimate, for a specified sample size and range of sample log SD, is represented by a point computed as the average of a very large number of simulated UCL values, since 10,000 UCL values were simulated for each type of UCL, sample size, and population CV. To make the plots more readable, the curves for each UCL on each plot were smoothed using locally estimated scatterplot smoothing (LOESS), a nonparametric smoothing spline technique developed for scatterplot smoothing (Fox and Weisberg, 2018). Another and more important reason to smooth the points in each curve is that the smoothed curves give an improved estimate of the expected value of each UCL loss function (risk of the UCL estimator) or of the expected value of the performance measure.

For each range of sample log SD, four figures are plotted as a set. The first figure in each set includes four plots that show the coverage, relative error variance, relative bias, and RelRMSE over a dense grid of sample sizes covering a wide range. The RelRMSE can be thought of as an accuracy measure that integrates relative bias and relative error variance into a single measure that is on the scale of average relative deviations from the true value. The second figure in each set is composed of three plots that show the risk of the UCLs with respect to the bounded loss function (Section 3.1) with parameters in Table 1; namely, " Loss_bnd_c," "Loss_bnd_m," and "Loss_bnd_a."

The third figure in each set has four plots that show the risk of the UCLs with respect to the unbounded loss function with LOR coverage loss (Section 3.2.1) and relative accuracy loss for the parameter values in Table 1. The fourth figure in each set is very similar to the third but shows the risk of the UCLs with respect to the unbounded loss function with probit coverage loss (Section 3.2.1) and relative accuracy loss for the parameter values in Table 1.

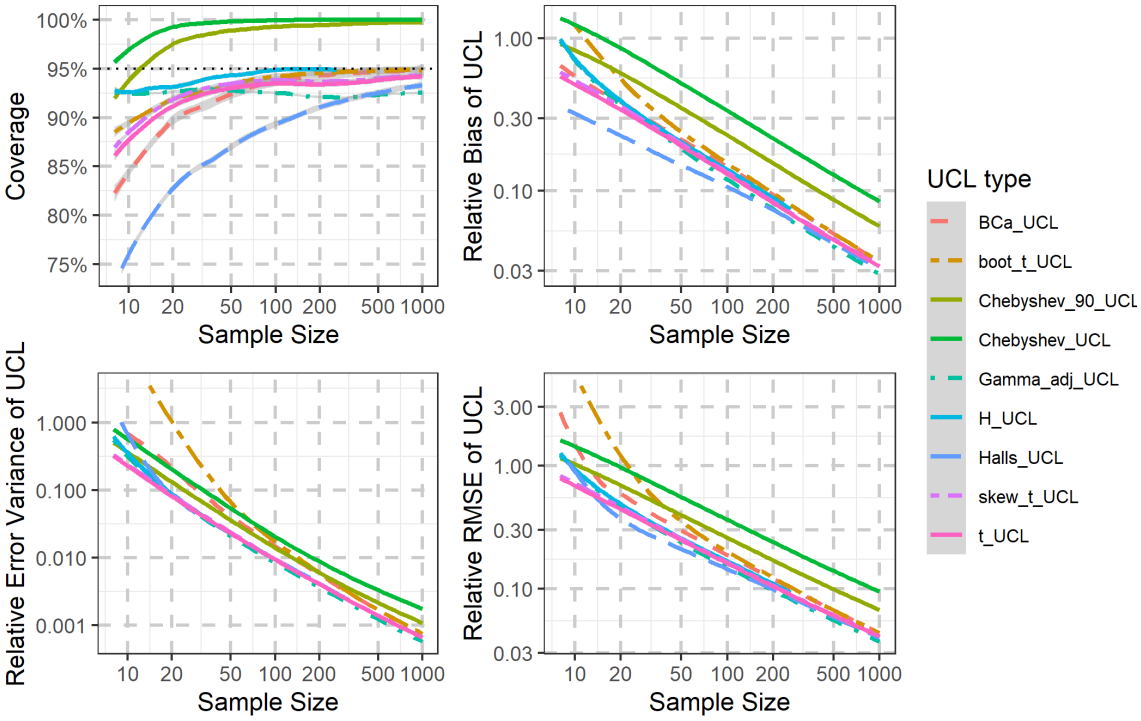

Figure 5. UCL summary for Lognormal with log SD in (0.0831,0.859]

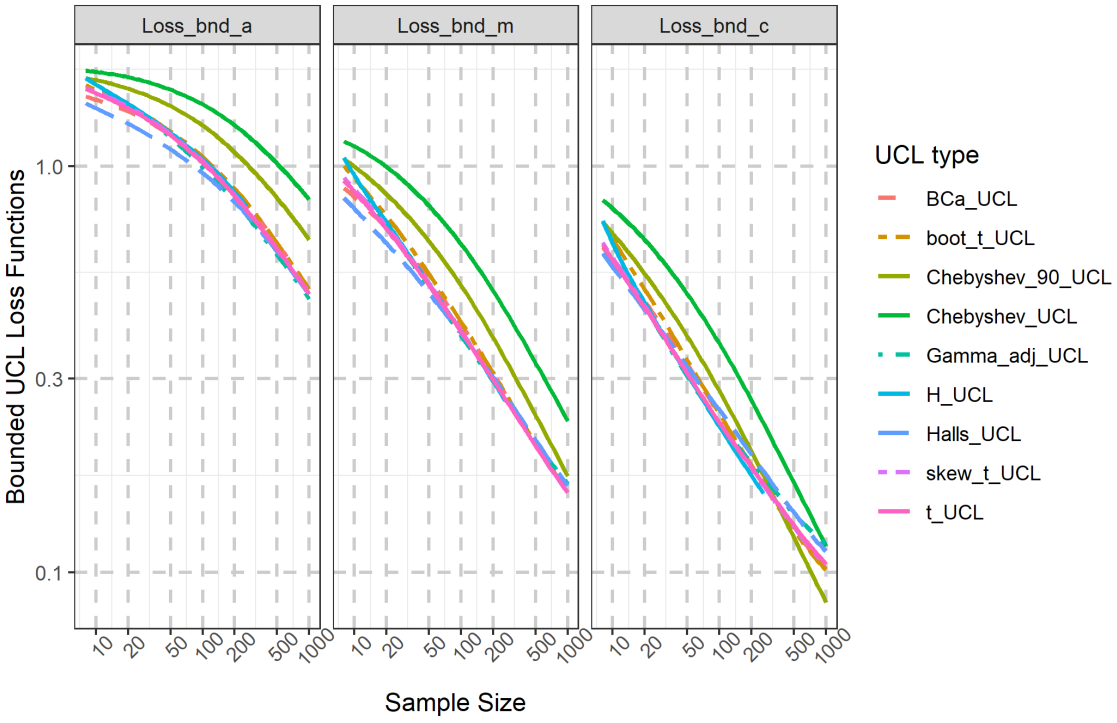

Figure 6. UCL Bounded Loss for Lognormal with log SD in (0.0831,0.859]

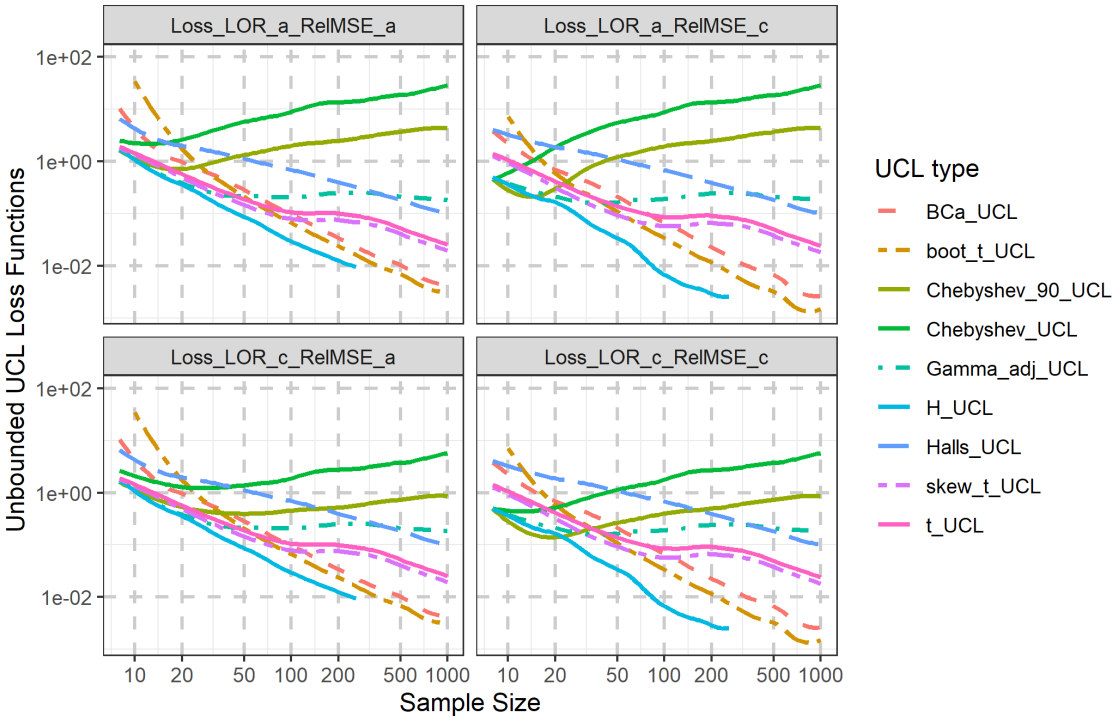

Figure 7. UCL Unbounded Loss with LOR loss for Coverage for Lognormal with log SD in (0.0831,0.859]

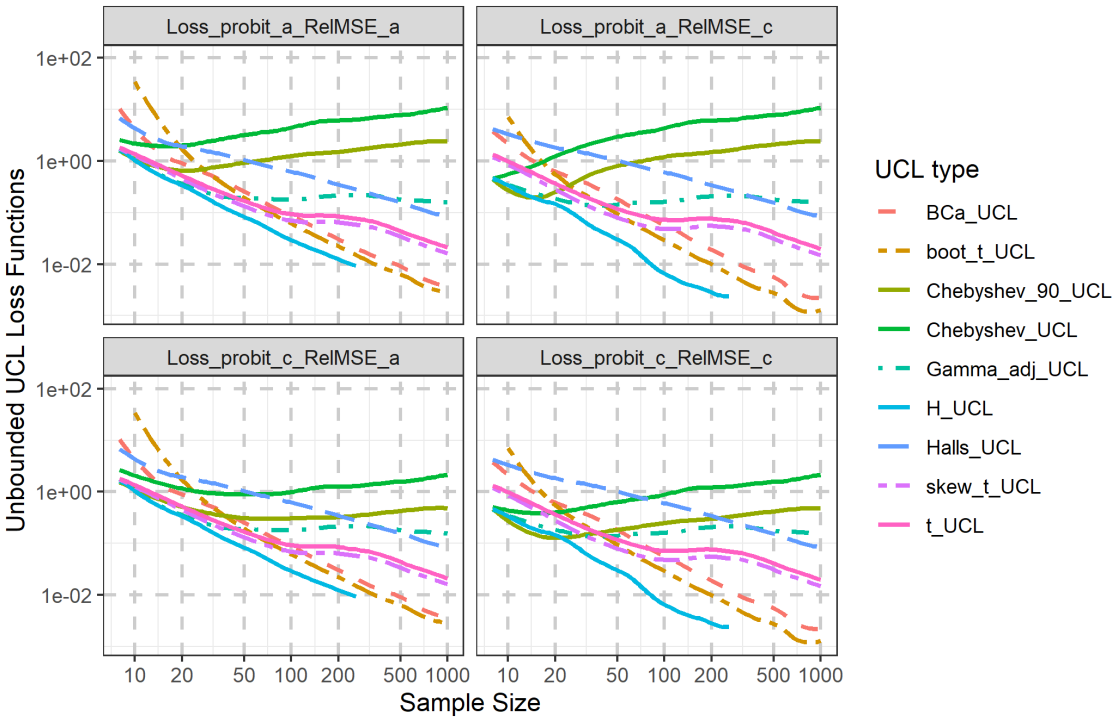

Figure 8. UCL Unbounded Loss with Probit loss for Coverage for Lognormal with log SD in (0.0831,0.859]

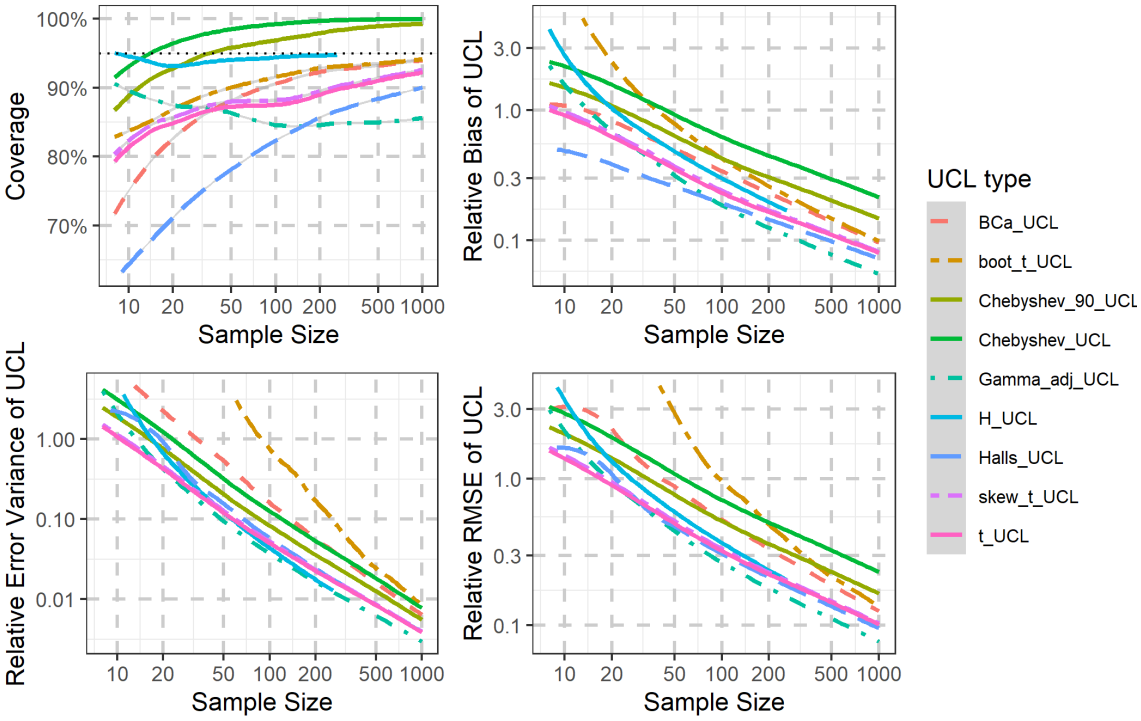

Figure 9. UCL summary for Lognormal with log SD in (0.859,1.37]

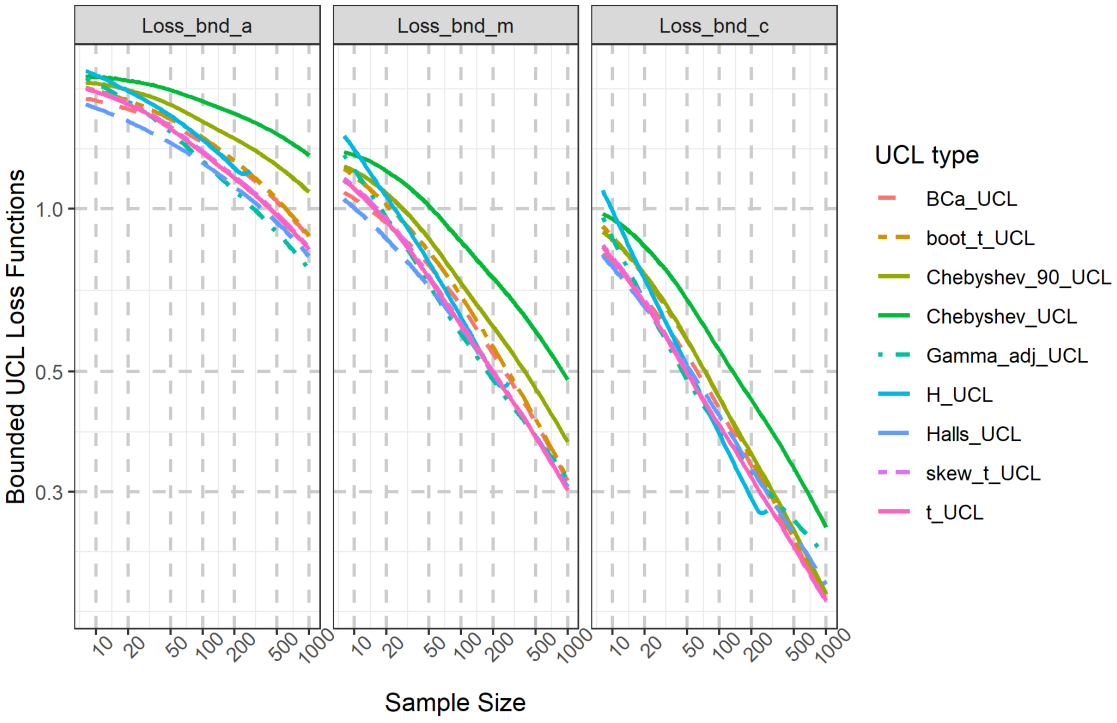

Figure 10. UCL Bounded Loss for Lognormal with log SD in (0.859,1.37]

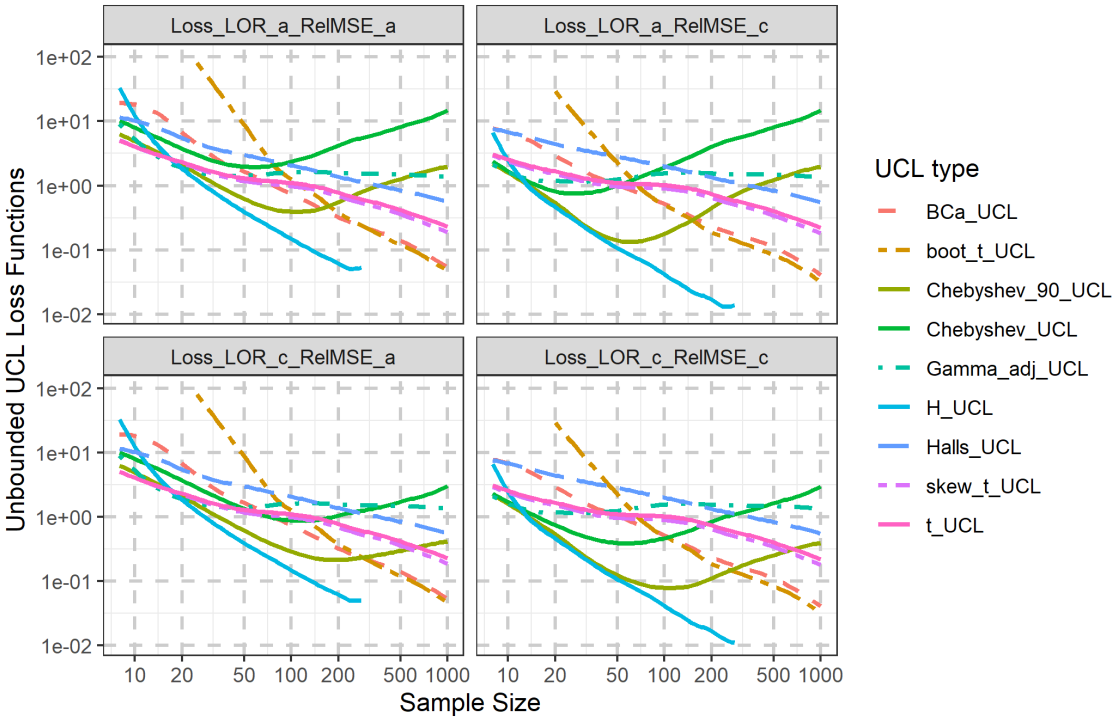

Figure 11. UCL Unbounded Loss with LOR loss for Coverage for Lognormal with log SD in (0.859,1.37]

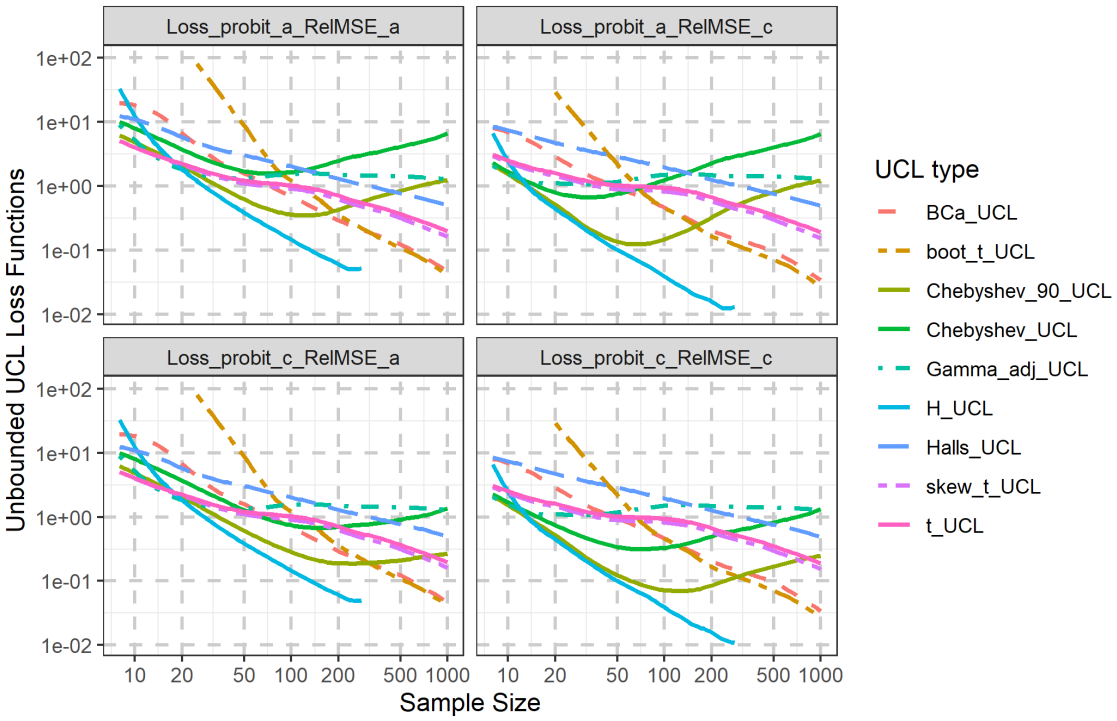

Figure 12. UCL Unbounded Loss with Probit loss for Coverage for Lognormal with log SD in (0.859,1.37]

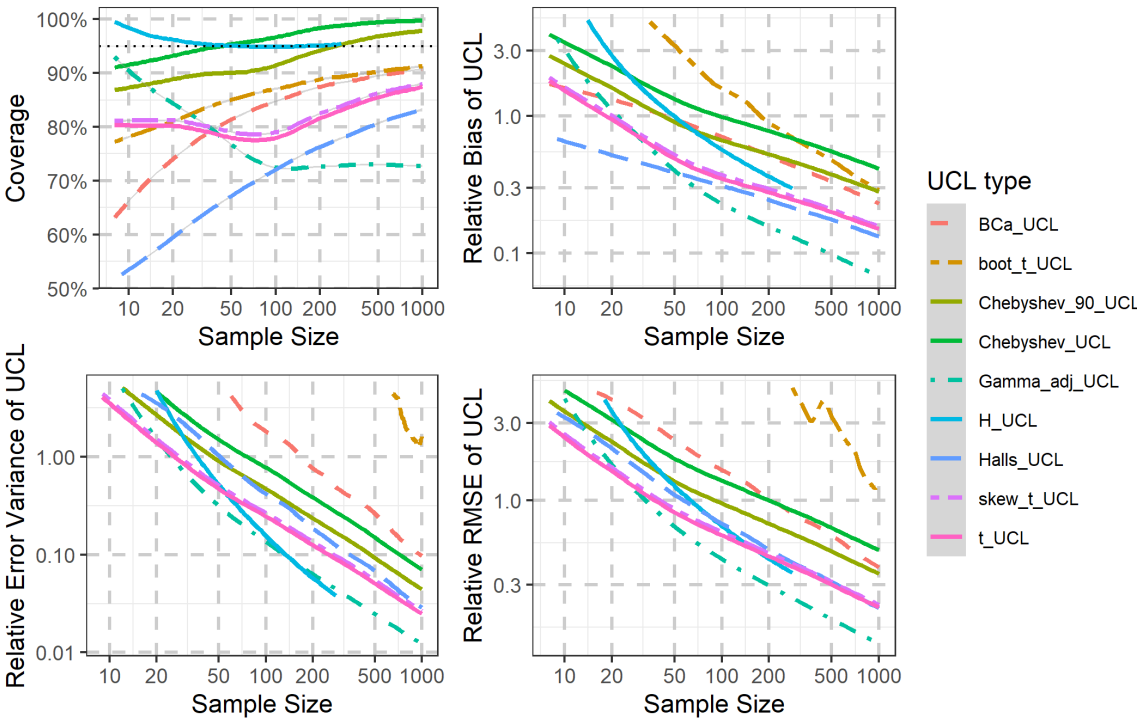

Figure 13. UCL summary for Lognormal with log SD in (1.37,1.81]

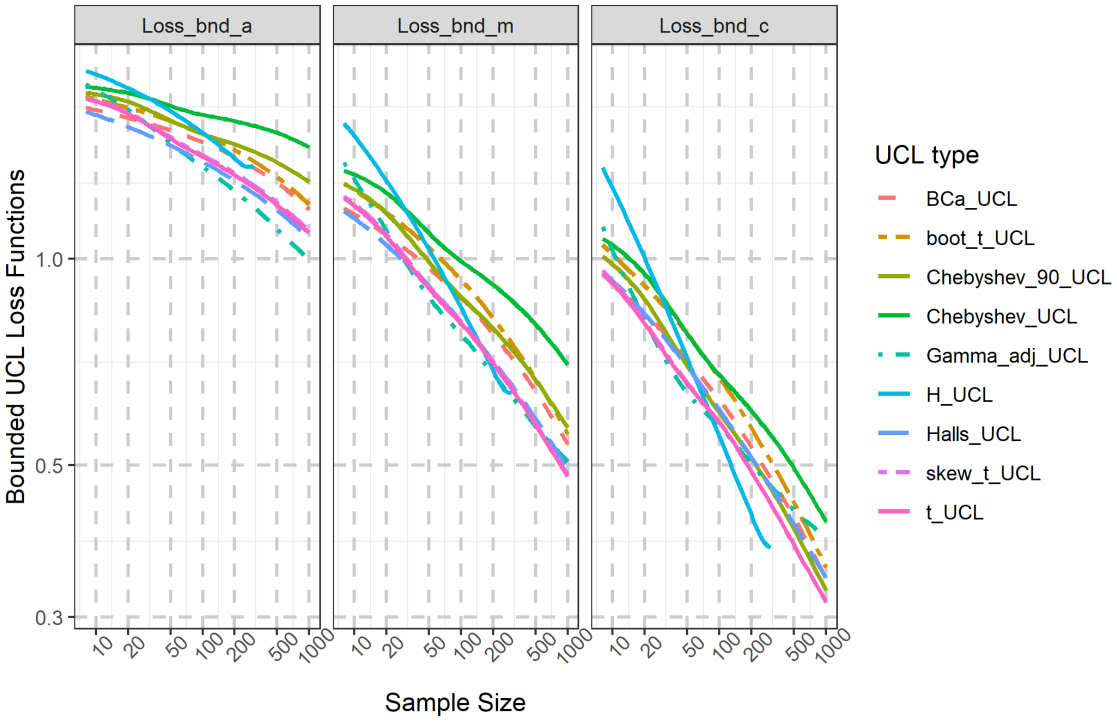

Figure 14. UCL Bounded Loss for Lognormal with log SD in (1.37,1.81]

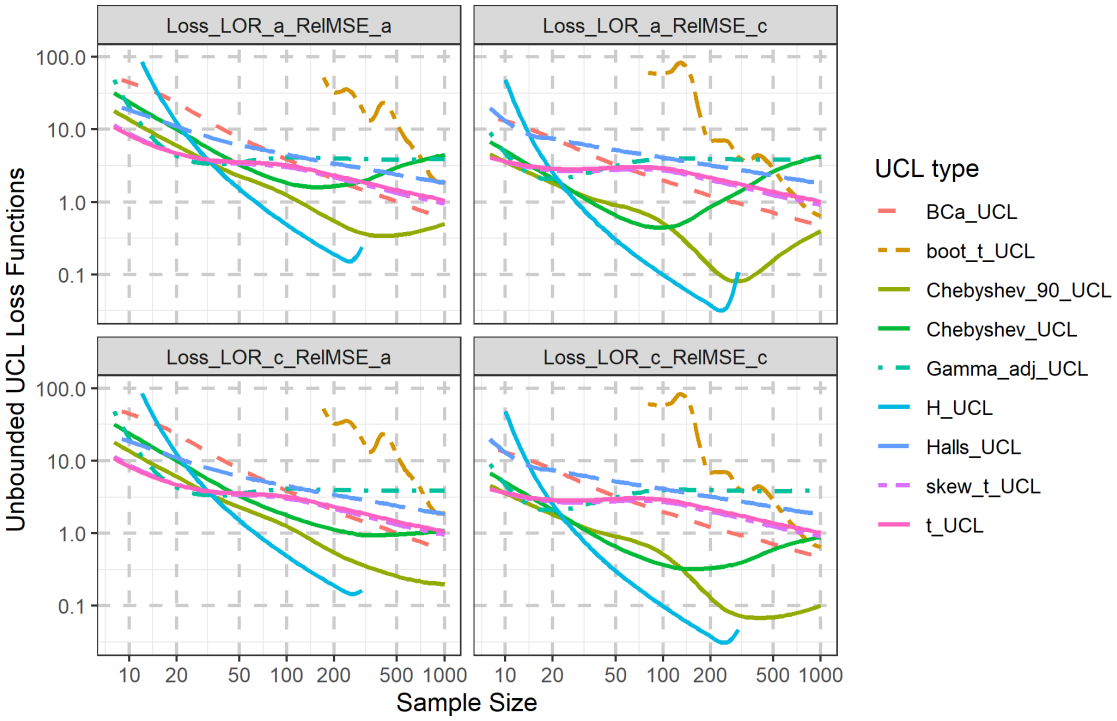

Figure 15. UCL Unbounded Loss with LOR loss for Coverage for Lognormal with log SD in (1.37,1.81]

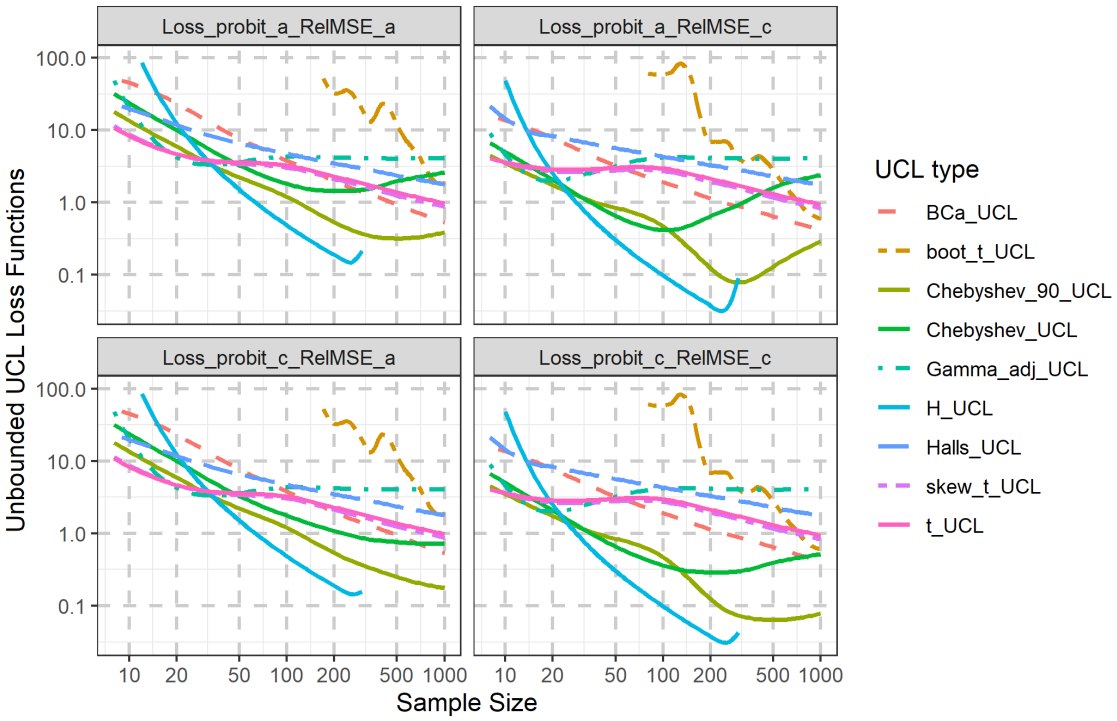

Figure 16. UCL Unbounded Loss with Probit loss for Coverage for Lognormal with log SD in (1.37,1.81]

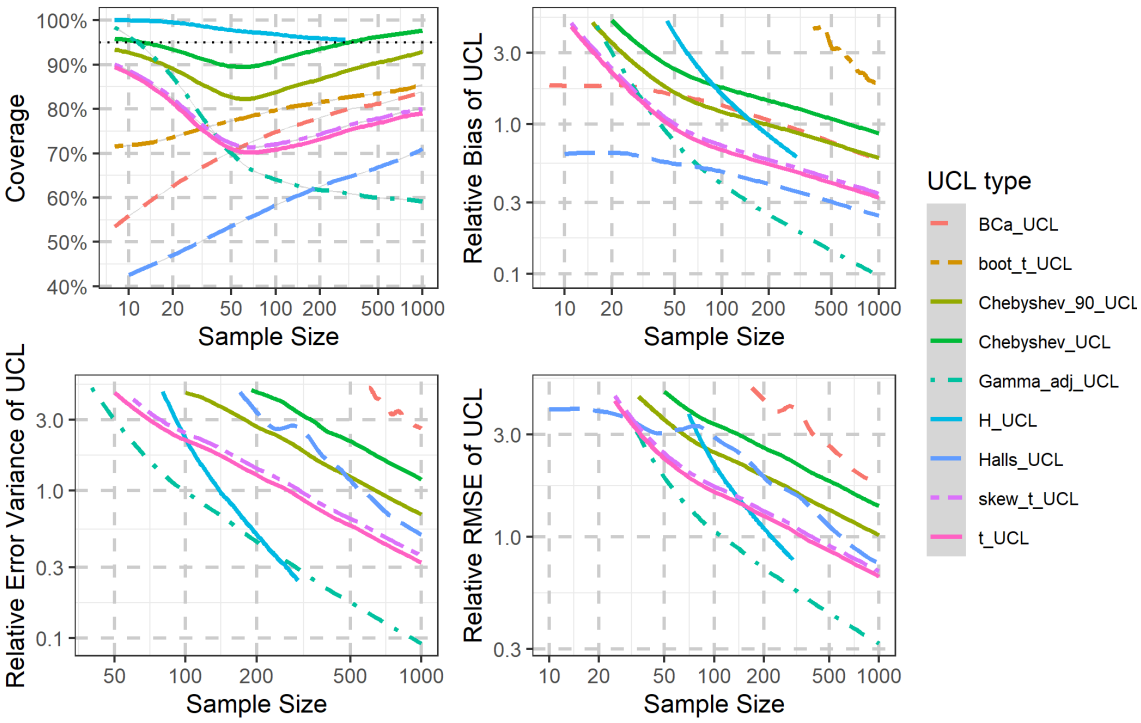

Figure 17. UCL summary for Lognormal with log SD in (1.81,5.41]

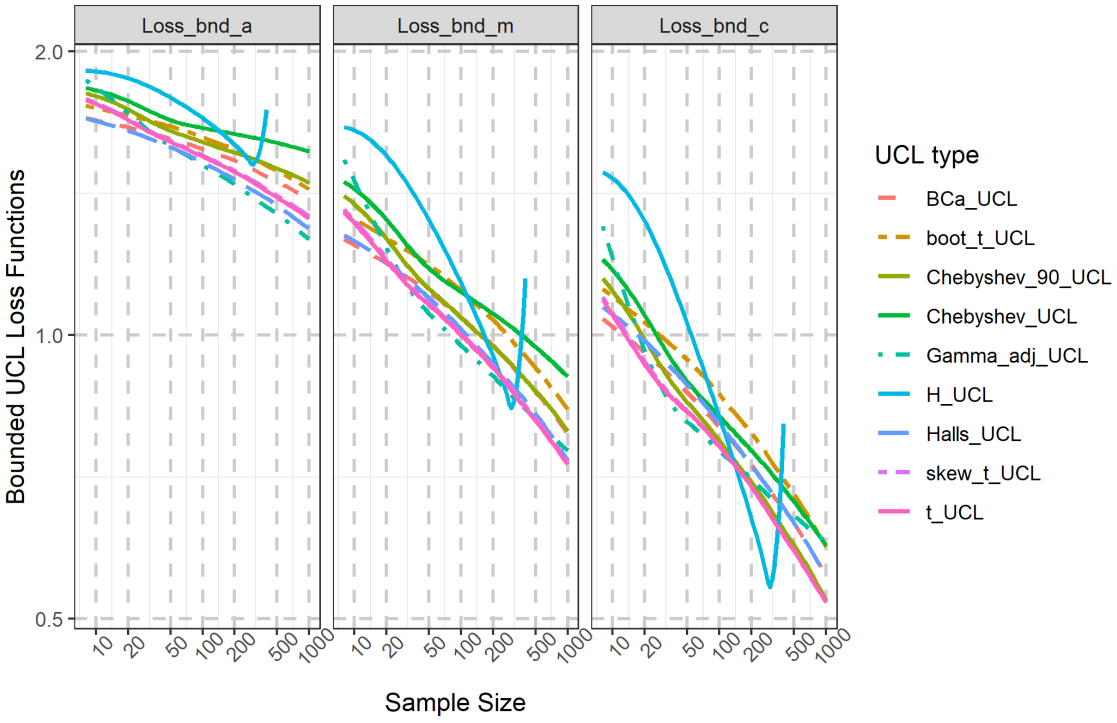

Figure 18. UCL Bounded Loss for Lognormal with log SD in (1.81,5.41]

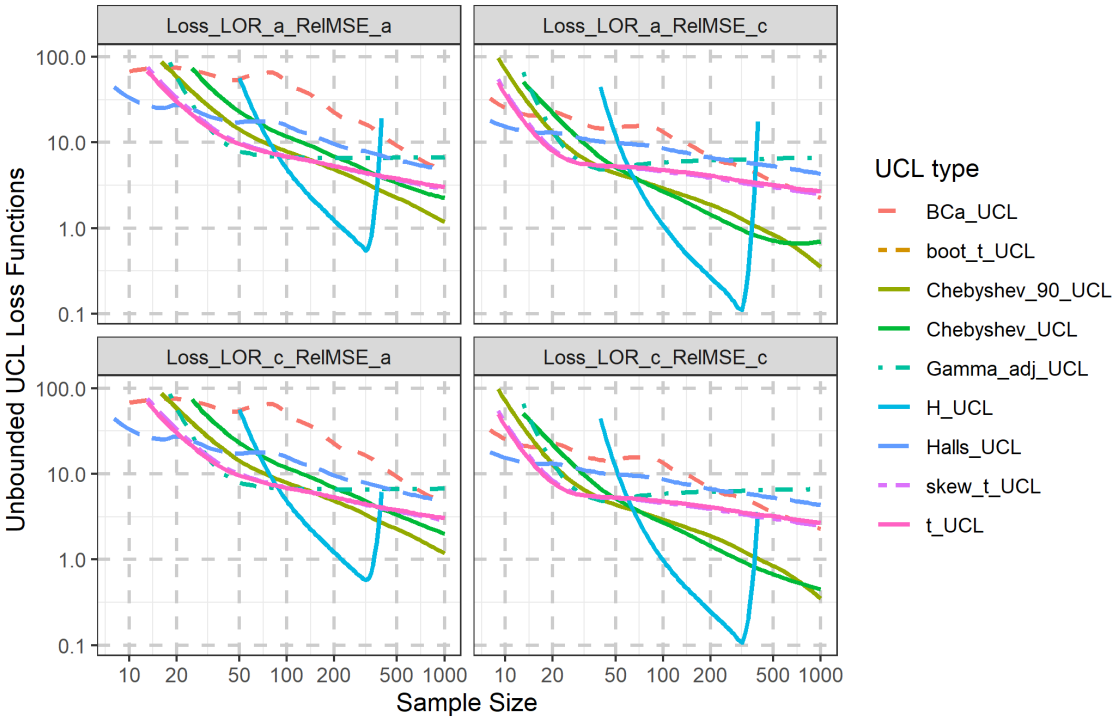

Figure 19. UCL Unbounded Loss with LOR loss for Coverage for Lognormal with log SD in (1.81,5.41]

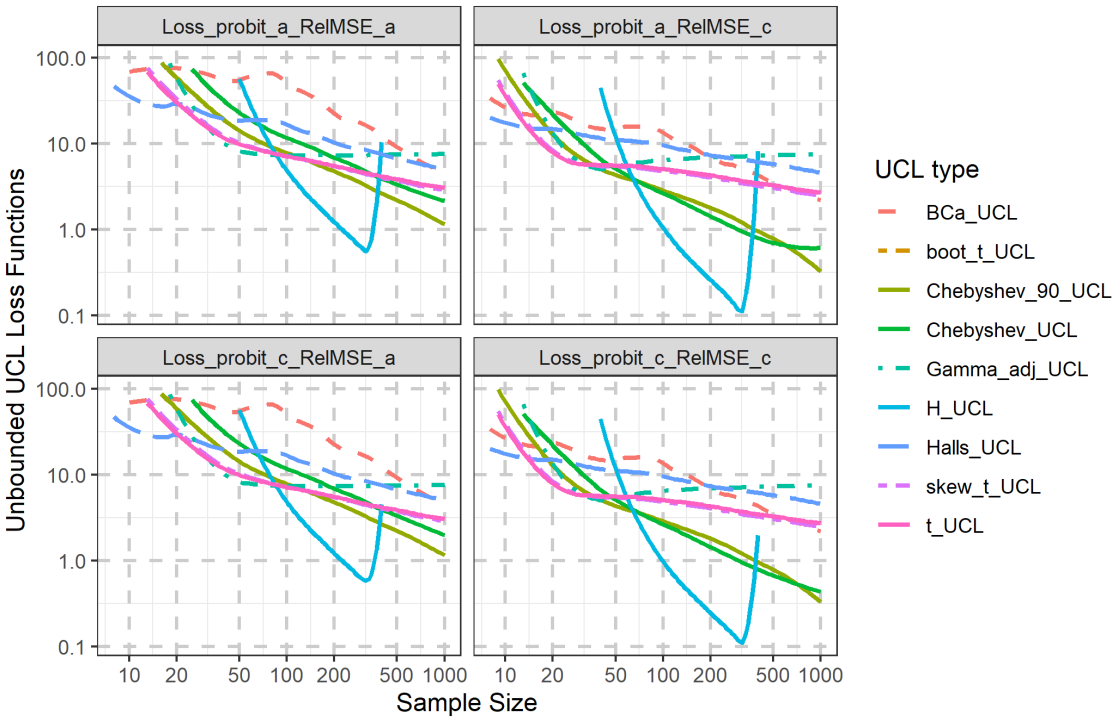

Figure 20. UCL Unbounded Loss with Probit loss for Coverage for Lognormal with log SD in (1.81,5.41]

4.1 Discussion of UCL Plots

As discussed in the preceding sections, the loss functions used were chosen to span a range of penalties for coverage error and inaccuracy. The purpose of comparing UCLs against a variety of relevant loss functions is to lend robustness to the conclusions of the comparisons.

Review of the plots above and of those in the Appendix A shows a number of very interesting things:

- The first is that the Chebyshev 90% and 95% UCLs perform quite poorly with respect to the accuracy measures relative bias and relative root MSE in the first figure in each set (Figures 5, 9, 13, and 17).
- Secondly, the Chebyshev 90% and 95% UCLs perform quite poorly with respect to the relative error variance measure in the first figure in each set (Figures 5, 9, 13, and 17). This is due to the fact that both of these multiply the sample standard deviation by relatively large factors.
- Thirdly, with respect to the varieties of bounded loss considered, the Chebyshev 90% and 95% perform overall the most poorly of any of the UCLs considered over all sample sizes and ranges of sample log SD, as seen in the second figure in each set (Figures 6, 10, 14, and 18).
- Fourthly, with respect to the varieties of unbounded loss considered, the Chebyshev 90% and 95% perform overall poorly compared to the H-UCL below sample size 250, as seen in the third and fourth figures in each set (Figures 7, 8, 11, 12, 15, 16, 19, and 20). In all the plots, the H-UCL behaves badly starting somewhere in the range of sample sizes 250 to 300. This is due to a numerical issue in the code in the EnvStats R library used to compute the H-UCL.
- Except for small sample sizes or very large sample log SD, the H-UCL is by far the closest to the target coverage (95%) of all the UCLs considered (Figures 5, 9, 13, and 17).
- Hall's bootstrap UCL has by far the lowest coverage and the lowest relative bias (although still biased significantly high) of all the UCLs, except in the case of large sample log SD combined with large sample size, in which case the Gamma adjusted UCL has lower coverage and less relative bias.
- The risk profiles of UCLs under the unbounded loss functions with LOR and probit coverage losses are similar but not identical.
- For sample sizes 20 and below falling into the largest category of sample log SD (≥ 1.8), Hall's bootstrap UCL has the lowest risk under all of the unbounded loss functions considered.
- The bootstrap-t UCL becomes wild for all sample sizes for sample log SD above 1.3.
- For both bounded and unbounded losses over the ranges of sample sizes and sample log SD considered, the risk performances of the t-UCL and the skewed t-UCL are very similar.

- For sample sizes of approximately 20–25 and smaller with sample log SD between approximately 0.8 and 1.8, the t-UCL has the lowest risk among the unbounded losses considered.
- The behavior of the UCLs appears to be better separated and characterized by the unbounded loss functions than by the bounded loss functions.

These simulated UCLs and summary statistics will be an important source of data moving forward to improve the lognormal UCL recommendation rules (that is, for data that has gone through the GoF logic described in Section 2.0 and is deemed to be lognormal) for ProUCL 5.2 and to substantially improve them in ProUCL 6.0. Since the patterns illustrated in the plots above are complex, decision rules for UCL recommendations can best be formulated with the help of machine learning methods.

5.0 UCL Recommendation Rules

ProUCL uses many different UCL methods for estimation, and, for many data sets, uses an underlying decision logic to recommend one of the estimated UCLs. This decision logic is based in part on the results of GoF tests but in the past has only addressed coverage of a UCL and not accuracy. For data sets that are considered approximately lognormal, it is common for ProUCL 5.1 to recommend a Chebyshev UCL. The objective of this simulation study and analysis is to improve these rules for data which ProUCL 5.2 has determined to be approximately lognormal based on the revised GoF rules described in Section 2.0.

5.1 Algorithm for Recommendation Rules

The new recommendation rules are determined using classification trees estimated using the method of Recursive Partitioning (RPart) as implemented in the R library rpart (version 4.1-15). The objective is to identify the UCL estimator that minimizes the aggregated risk measures for various values of decision variables. The decision variables are the sample size (N) and the log SD of the samples in each simulated data set.

The aggregated risk measures used are derived from the eight unbounded loss functions used in the simulation study. The bounded loss functions are useful but not as informative as the unbounded loss functions, since the bounded loss functions don't separate the UCL estimators as well. Two types of aggregated risk measures are used. The first computes the average value across loss functions for each combination of UCL type, N, and log SD category over all the filtered simulated data sets. The second computes the maximum across the loss functions for each combination of UCL type, N, and log SD category over all the filtered simulated data sets. Since the values of log SD are continuous, in order to match the values up, the values of log SD are binned into 100 categories using the min, max, and percentiles of the log SD value in the simulated data. The use of several different loss functions, some more conservative (emphasizing coverage and allowing more overestimation) and some more accurate (less extreme overestimation but not necessarily as good coverage), to derive the aggregated risk measures makes the conclusions derived from the study more robust.

For each type of aggregated risk and for each combination of values of N and log SD category, the UCL estimator with the smallest risk is identified. The selected UCLs are assigned to UCL_{MAL} (UCL type with minimum average loss) and UCL_{MML} (UCL type with minimax loss) which take a value (UCL name) for each combination of values of N and log SD category. The former is the choice of UCL type that would perform best on average for a given value of N and sample log SD. The latter is a minimax choice of UCL type for a given value of N and sample log SD.

Since the values of the average loss for each variety of loss function are estimates based on simulation, so also are the risk values and the selected UCL types for UCL_{MAL} and UCL_{MML}. The assignments of UCL_{MAL} and UCL_{MML} for the combinations of N and log SD do potentially contain some errors. Therefore, a statistical classification method must be used to extract the signal from the noise and to develop relatively simple recommendation rules that can be easily implemented.

5.2 Risk Profiles

Figure 21 and Figure 22 below illustrate the features of the aggregated risk measures across ranges of the sample log SD and for a range of sample sizes. These two figures effectively summarize the eight unbounded loss figures in Section 4.0 and the 20 unbounded loss figures in Appendix A. While the average risks are lower than the maximum risk levels, the patterns in these plots are very similar.

Figure 21 shows the average risk profiles for the various UCL estimators averaged across the various unbounded loss functions.

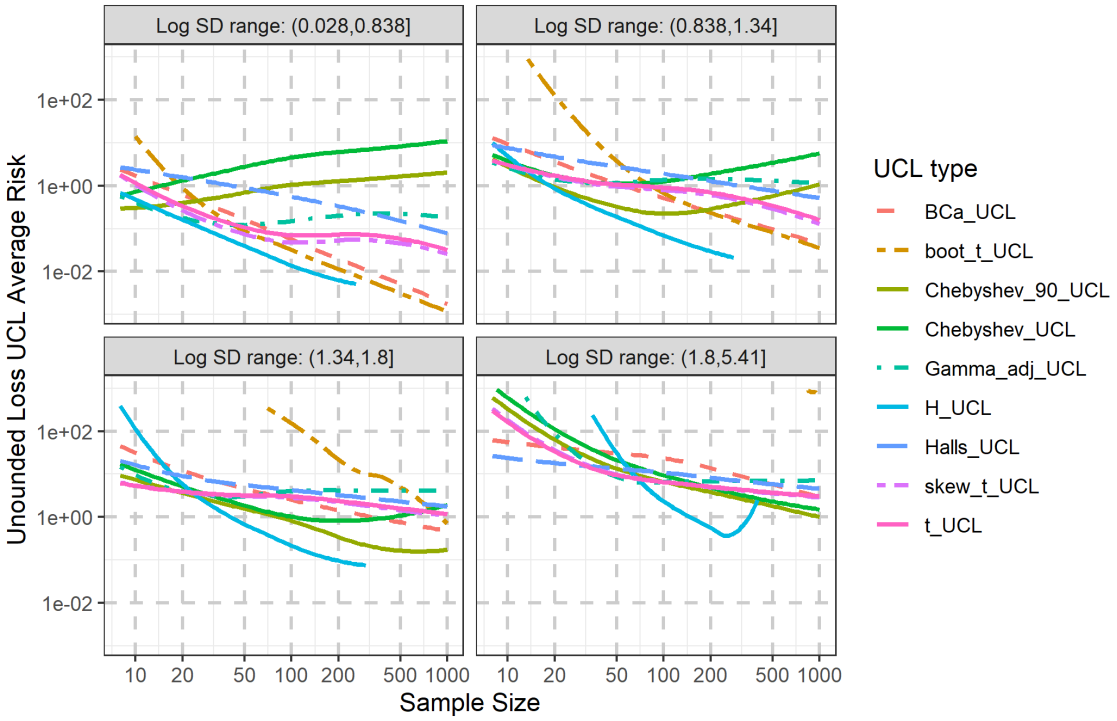

Figure 21. Average of eight unbounded loss functions for various UCLs by log-scale SD and sample size

Figure 22 shows the maximum risk profiles for the various UCL estimators averaged across the various unbounded loss functions.

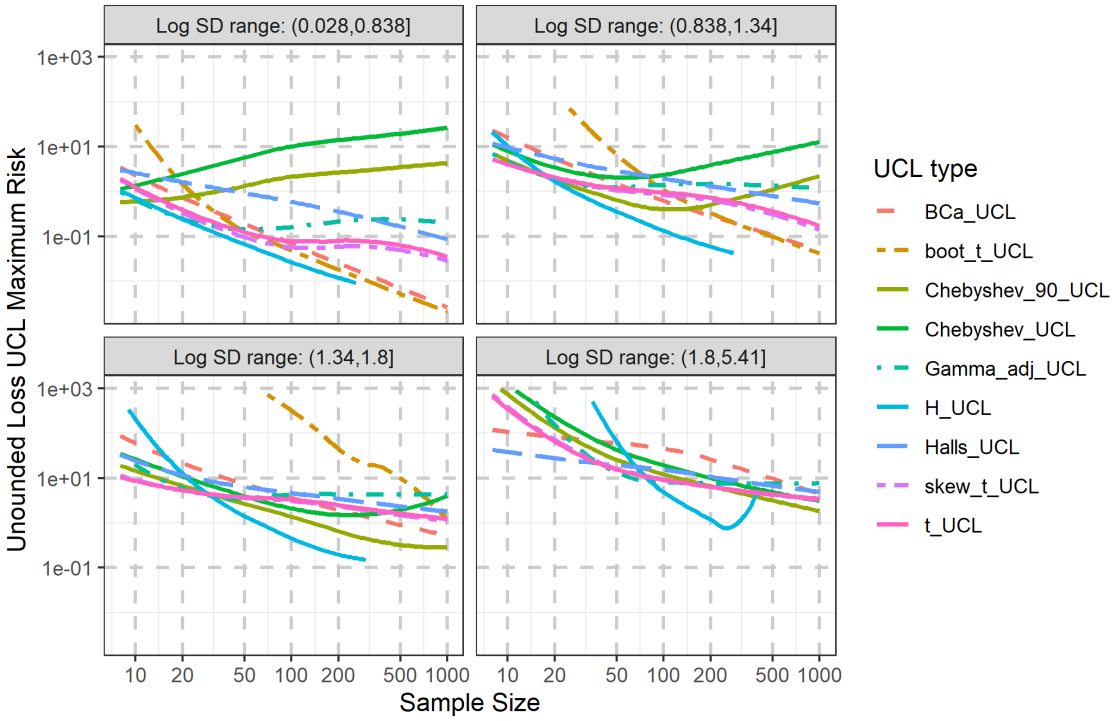

Figure 22. Maximum of eight unbounded loss functions for various UCLs by log-scale SD and sample size

These plots highlight features of the UCLs and suggest that the only real contenders for recommendation are the H-UCL in most cases, the t-UCL for small sample sizes combined with medium to large sample log SD, and possibly Hall's bootstrap UCL for small sample size combined with very large sample log SD. While the figures above are instructive, we proceed next to quantitative modeling in order to create recommendation rules.

5.3 UCL Recommendation Modeling

For deciding on the best recommendations for UCLs, the decision tree approach, Recursive Partitioning (RPart), is used. This is based on the original algorithm developed by Friedman (1977) as a Bayesian nonparametric classifier and as further developed and implemented by Therneau and Atkinson (2019) in the rpart R package. RPart is a classification tree algorithm that takes data with labels and predictor variables and performs successive binary splits of the data based on values of the predictors in order to classify the data into sets that predict the labels as accurately as possible. For these models, the prior distribution on UCL types is uniform; that is, all of the UCL types are treated equally. Also, the decision trees are built not based on the aggregated risks but rather based on the numbers of cases in which a particular UCL type is designated UCL_{MAL} or UCL_{MML} as a function of N and sample log SD. The aggregated risk numbers are already built into the UCL_{MAL} and UCL_{MML} designations.

To avoid the numerical problems with computation of the H-UCL for large sample size, the data used for computing the RPart trees uses only sample sizes of 250 and less. It is believed that, for sample sizes above 250, if the H-UCL is computed accurately, it will still have similar risk performance relative to the other UCL types as it does at sample size 250.

For estimating recommendation rules based on UCL_{MAL}, it is desirable to use a relatively simple rule with no more than one split in N and one in log SD. The tree is therefore pruned until it has only two levels. The tree (Figure 23) and associated decision rules for minimizing the average risk (Table 2) are shown below. The UCL type shown at each node in the tree is the one most likely to minimize the average risk, conditional on the predictor values being in the indicated ranges.

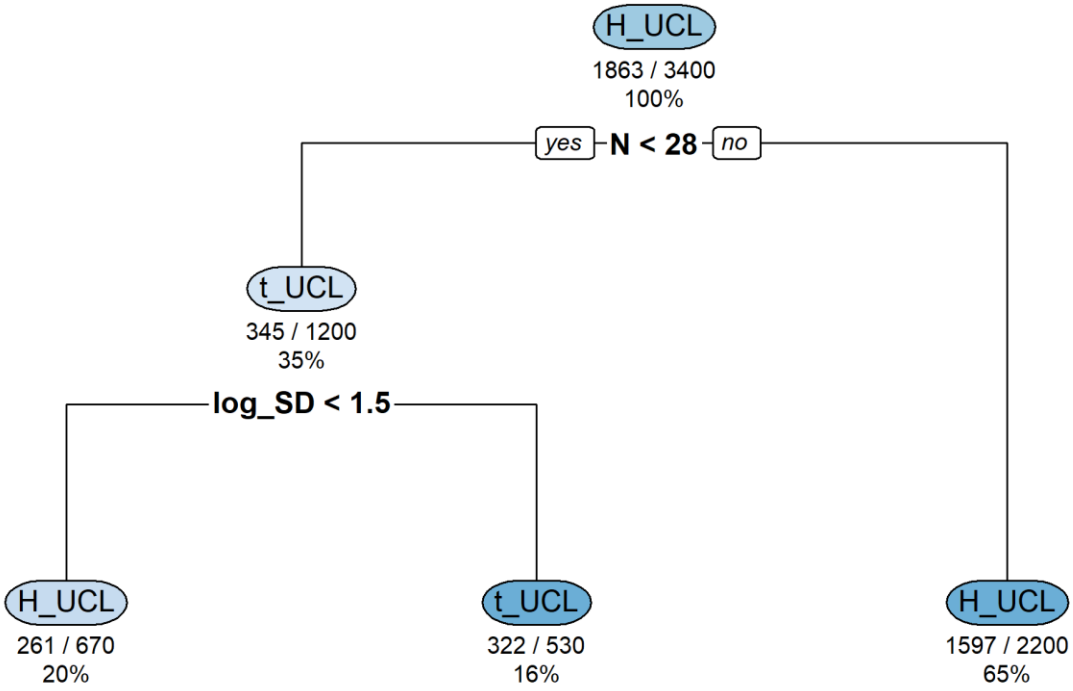

Figure 23. Decision tree for minimum average loss, pruned to two levels

Only the H-UCL and the t-UCL are recommended. The breakpoint in N is 28, and for smaller samples the t-UCL is recommended for cases in which the log SD is greater than or equal to 1.5.

Table 2. Decision rule output for UCL minimum average risk tree

```
       ske Gam Che Che H_U t_U Hal boo BCa
H_UCL [.06 .25 .22 .01 .39 .03 .00 .03 .01] when N <  28 & log_SD <  1.5
H_UCL [.03 .08 .03 .00 .73 .02 .01 .07 .03] when N >= 28
t_UCL [.04 .19 .00 .00 .01 .61 .15 .00 .00] when N <  28 & log_SD >= 1.5
```

The decision tree for minimax risk, when pruned to two levels, gives a very similar result. The only difference is that the decision point for sample log SD is 1.3 instead of 1.5.

5.4 Tentative Lognormal UCL Recommendations

It is reasonable to compromise between the minimum average risk rule and the minimax risk rule. Therefore, the rule that for $N \geq 28$ use the H-UCL and that for $N < 28$, sample log SD < 1.4, use the H-UCL and otherwise the t-UCL can be tentatively recommended.

It must be strongly pointed out again that these recommendations are for data generated from a lognormal distribution and that have passed through the GoF screening procedure described in Section 2.0. Passing the GoF screening procedure means that ProUCL 5.2 is treating the data as lognormal. Some of the lognormal data, especially that simulated with the smallest CV (0.01), was screened out as normal (was not rejected by either test of normality at level 0.1). Some of the lognormal data simulated with small to moderate CV was screened out as Gamma (was not rejected by either Gamma GoF test at level 0.05). After these two filters, the remaining data was accepted as lognormal only if it was not rejected by two tests of lognormality at level 0.1.

Although this recommendation is for data treated by ProUCL 5.2 as lognormal, it would be appropriate to run further simulations with data from other right-skewed distributions, including mixture distributions, to confirm these recommendations. It must also be pointed out that the effects of detection limit censoring were not modeled in this simulation. It is suspected that data sets with a large number of nondetects could create problems for the H-UCL. This should also be explored.

6.0 Conclusion

The results of this study provide clear and convincing evidence that neither the Chebyshev 95% UCL nor the Chebyshev 90% UCL are useful procedures for constructing UCLs for data deemed to be lognormal.

Furthermore, analysis of the lognormal UCL simulation study data using RPart classification trees for risk minimization allows formulation of a simple tentative recommendation rule for UCLs in data classified as lognormal by ProUCL 5.2. That rule may be simply stated as:

H-UCL when $N \geq 28$ or log-scale SD ≤ 1.4, and the t-UCL otherwise.

It is important that this recommendation for lognormal data be confirmed by simulations with other right-skewed distributions and by accounting for the effects of detection limit censoring.

7.0 References

Berger, J. O. (1985). *Statistical Decision Theory and Bayesian Analysis, 2nd edition*. Springer-Verlag.

Bickel, P.J. and K.A. Doksum (1977). *Mathematical Statistics: Basic Ideas and Selected Topics*. Holden-Day. San Francisco.

Casella, G., Hwang, J.T., and Robert, C. (1990). Loss Functions for Set Estimation. Biometrics Unit Technical Report BU-999-M, Cornell University.

Casella, G. and Hwang, J. T. (1991). EVALUATING CONFIDENCE SETS USING LOSS FUNCTIONS. *Statistica Sinica*, 1(1), 159–173. URL:http://www.jstor.org/stable/24303998

Casella, G., Hwang, J. T. G., & Robert, C. (1993). A PARADOX IN DECISION-THEORETIC INTERVAL ESTIMATION. *Statistica Sinica*, 3(1), 141–155. URL:http://www.jstor.org/stable/24304942

Davison, A.C. and D.V. Hinkley (1997). *Bootstrap methods and their application*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press.

Fox, J. and S. Weisberg (2018). "Appendix: Nonparametric Regression in R" (PDF). An R Companion to Applied Regression (3rd ed.). SAGE. ISBN 978-1-5443-3645-9.

Friedman, J. H. (1977). A recursive partitioning decision rule for nonparametric classification. *IEEE Trans. Computers*, *26*(4), 404-408.

Hollander, M. and D.A. Wolfe (1973). *Nonparametric Statistical Methods*. John Wiley & Sons.

Lehmann, E.L. (1983). *Theory of Point Estimation.* Wadsworth & Brooks. Pacific Grove, CA.

Lehmann, E.L. (1986). *Testing Statistical Hypotheses: Second Edition.* John Wiley & Sons.

Meeden, G. and Vardeman, S. (1985). Bayes and admissible set estimation. *J. Amer. Statist. Assoc.* 80, 465-471.

Mood, A.M., F.A. Graybill, and C.D. Boes (1974). *Introduction to the Theory of Statistics*. McGraw-Hill.

Randles, R.H. and D.A. Wolfe (1991). *Introduction to the Theory of Nonparametric Statistics*. Krieger Publishing Co., Malabar, Florida. 450 pp.

Rao, C.R. (2002). *Linear Statistical Inference and Its Applications: 2nd Edition*. John Wiley and Sons.

Serfling, R. (1980). *Approximation Theorems of Mathematical Statistics.* John Wiley and Sons.

Terry Therneau and Beth Atkinson (2019). rpart: Recursive Partitioning and Regression Trees. R package version 4.1-15. https://CRAN.R-project.org/package=rpart

Winkler, R. L. (1972). A Decision-Theoretic Approach to Interval Estimation. *J. Amer. Statist. Assoc.* 61, 187-191.

Appendix A: Detailed UCL Plots using Deciles of Log SD

The following plots are the same as in Section 4.0, except that there the UCL summary statistics and loss function values are plotted in groups by quartile of the log SD of the data sets from which the UCLs were computed. Here the data are grouped for plotting by decile of the log SD of the data sets. This gives a more detailed view of the behavior of the UCLs.

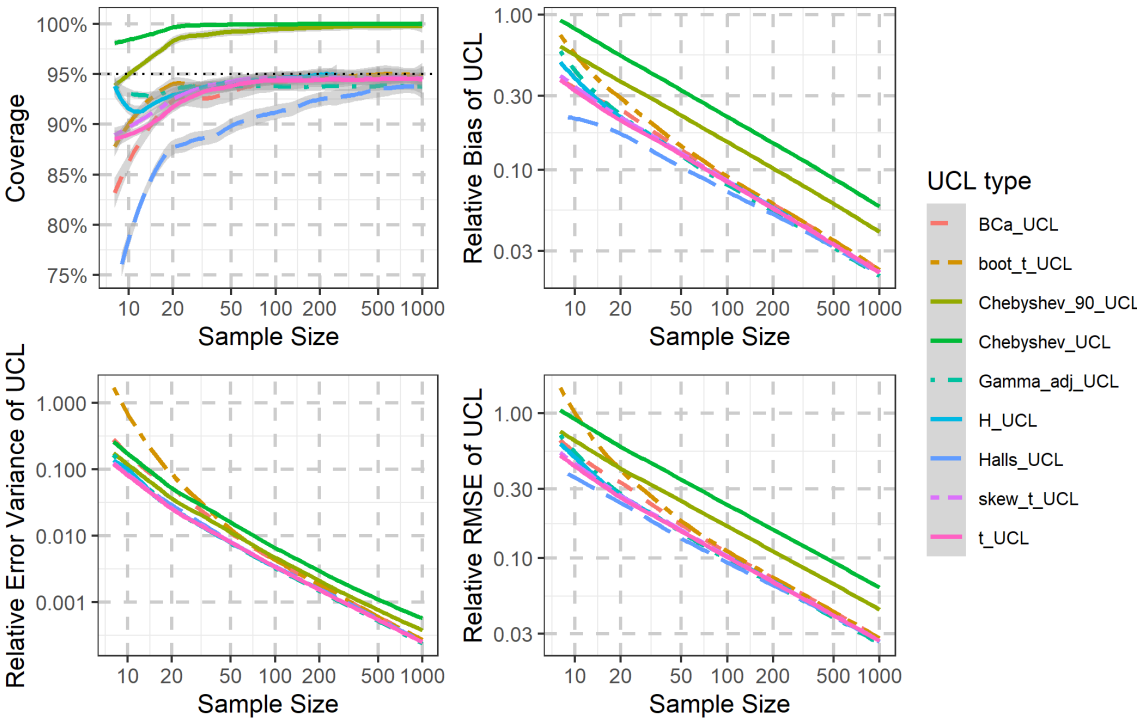

Figure 24. UCL summary for Lognormal with Std Dev of Logs in (0.0831,0.576]

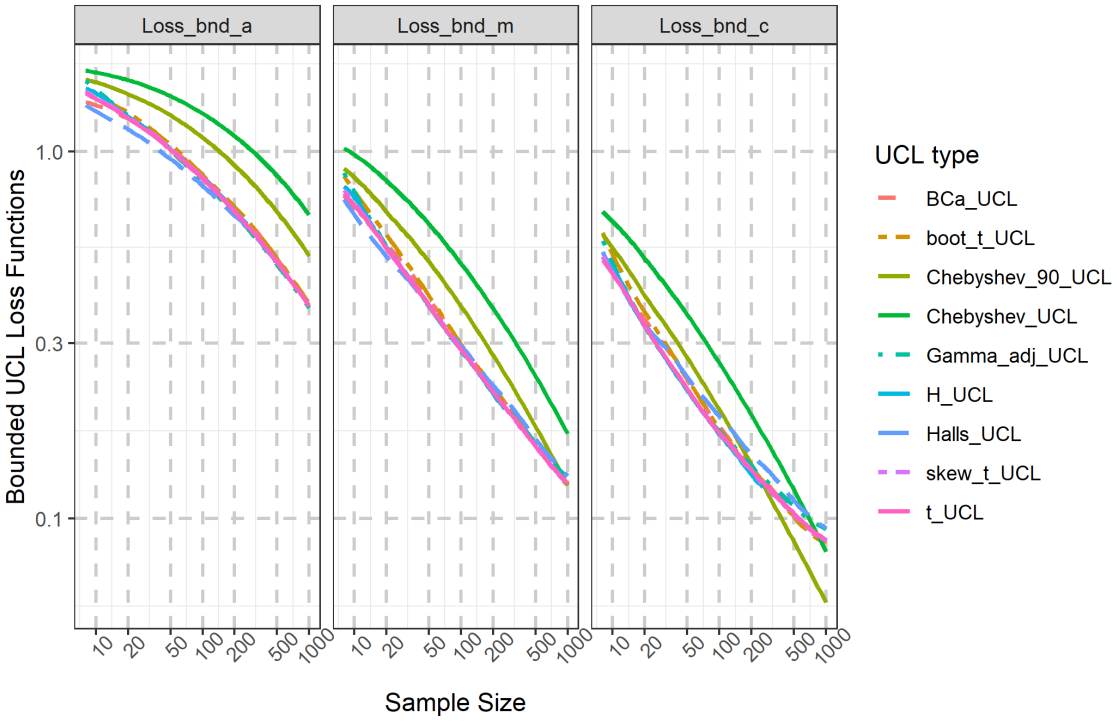

Figure 25. UCL Bounded Loss for Lognormal with Std Dev of Logs in (0.0831,0.576]

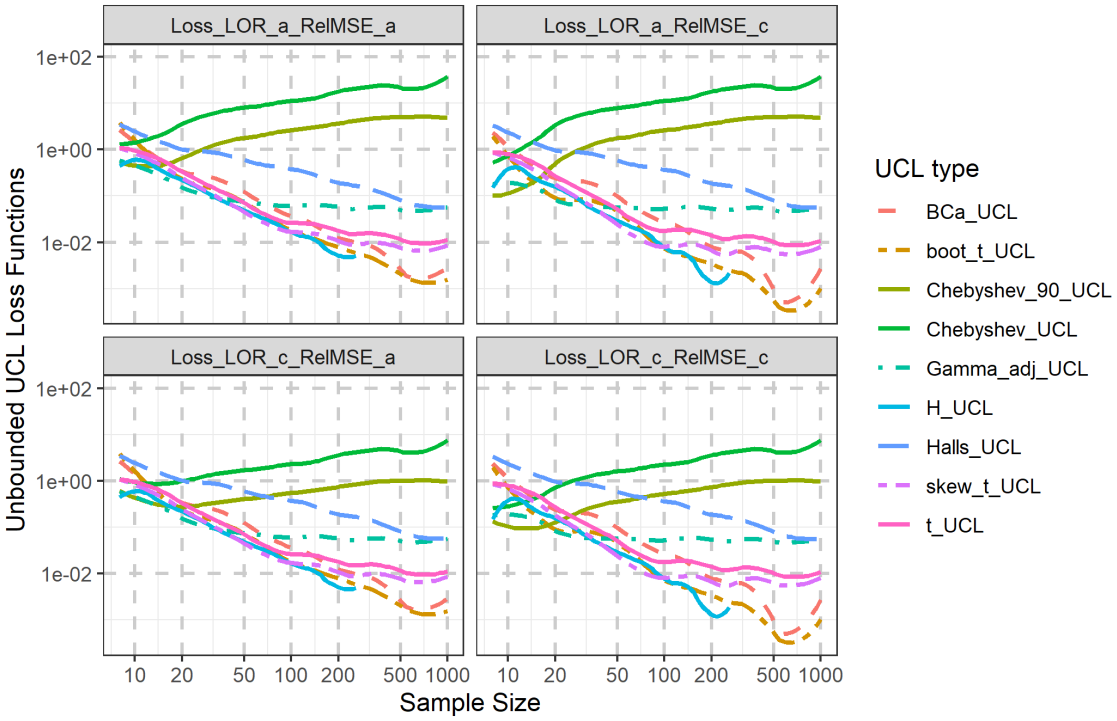

Figure 26. UCL Unbounded Loss with LOR loss for Coverage for Lognormal with Std Dev of Logs in (0.0831,0.576]

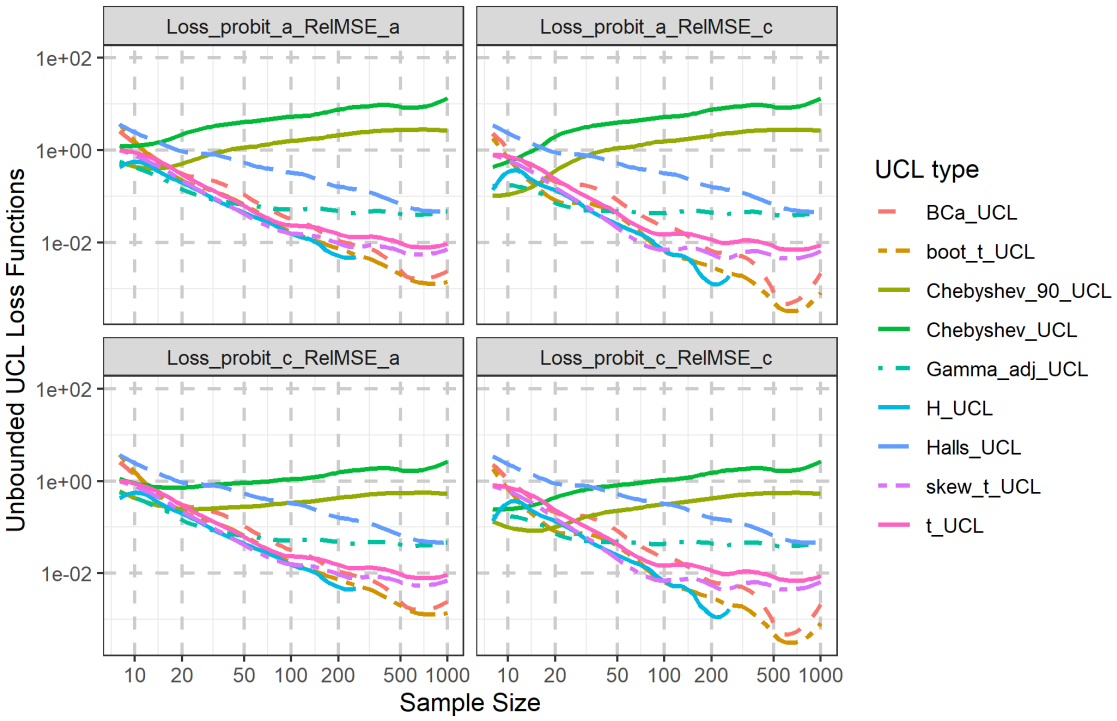

Figure 27. UCL Unbounded Loss with Probit loss for Coverage for Lognormal with Std Dev of Logs in (0.0831,0.576]

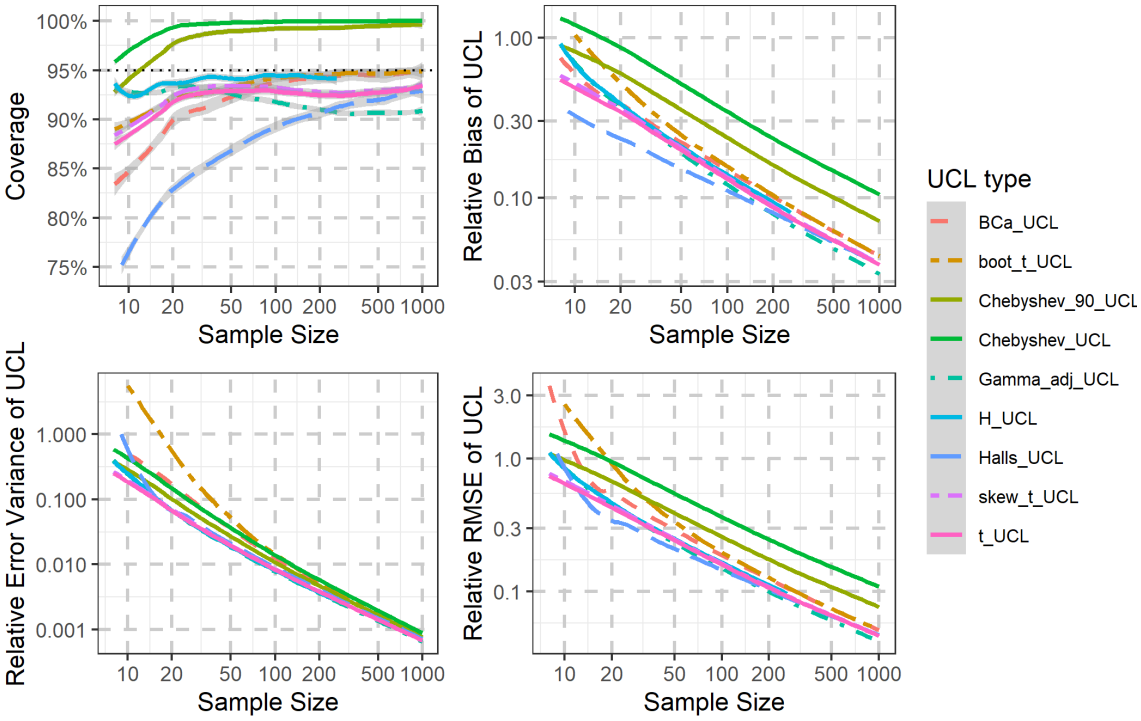

Figure 28. UCL summary for Lognormal with Std Dev of Logs in (0.576,0.773]

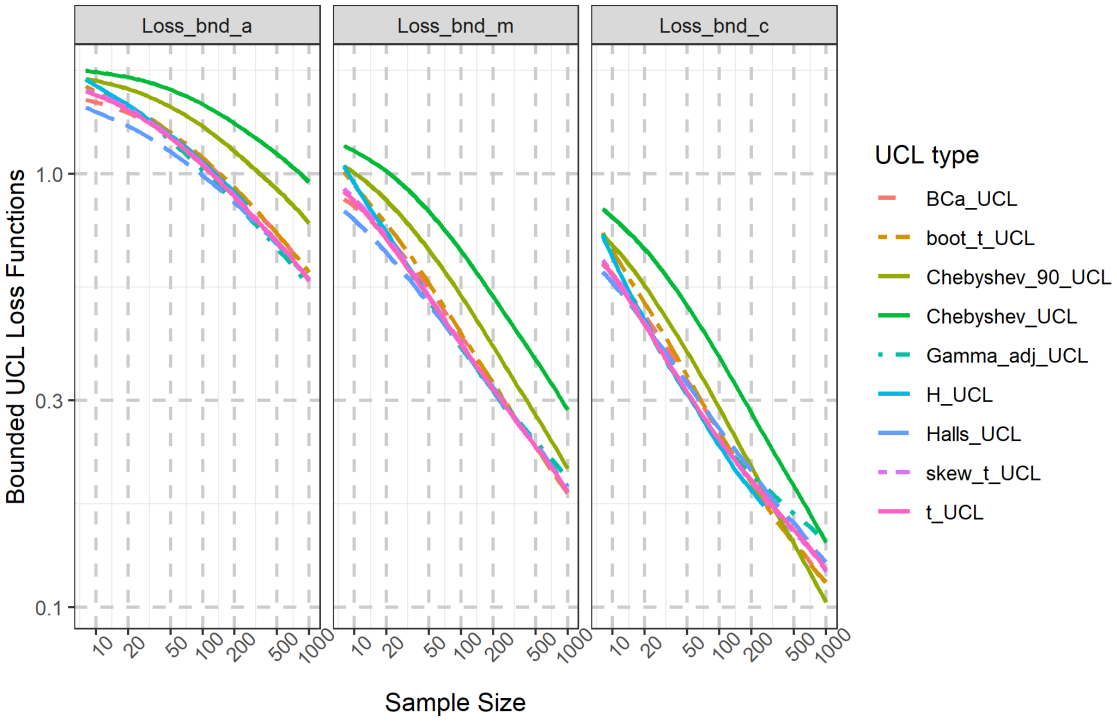

Figure 29. UCL Bounded Loss for Lognormal with Std Dev of Logs in (0.576,0.773]

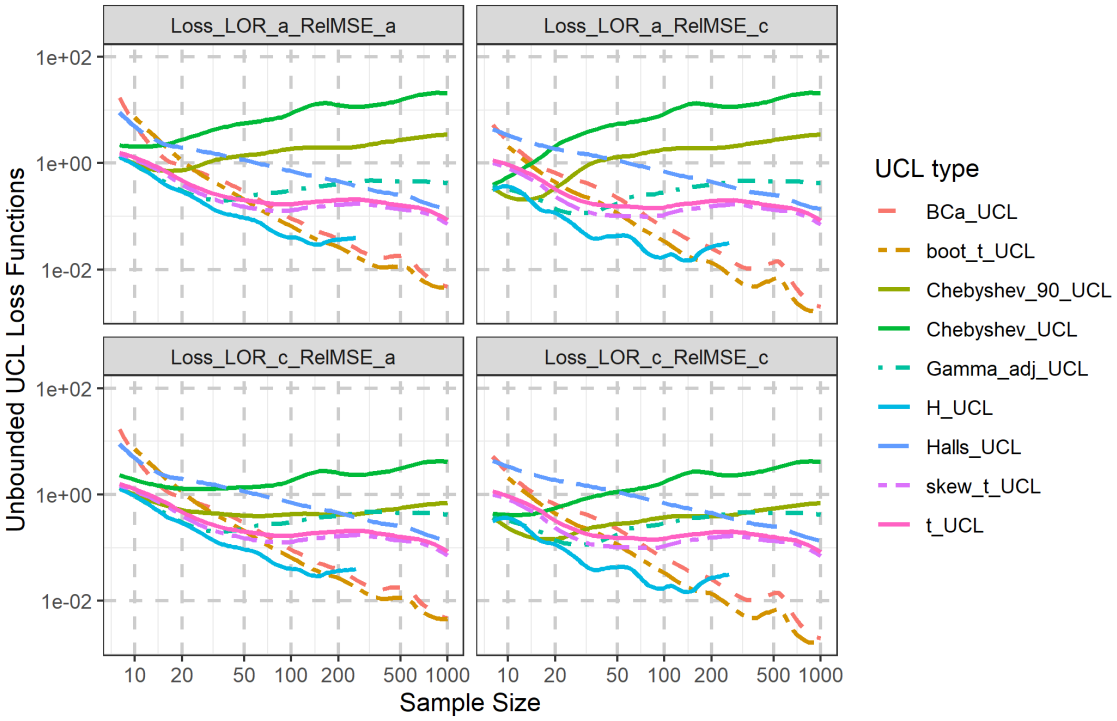

Figure 30. UCL Unbounded Loss with LOR loss for Coverage for Lognormal with Std Dev of Logs in (0.576,0.773]

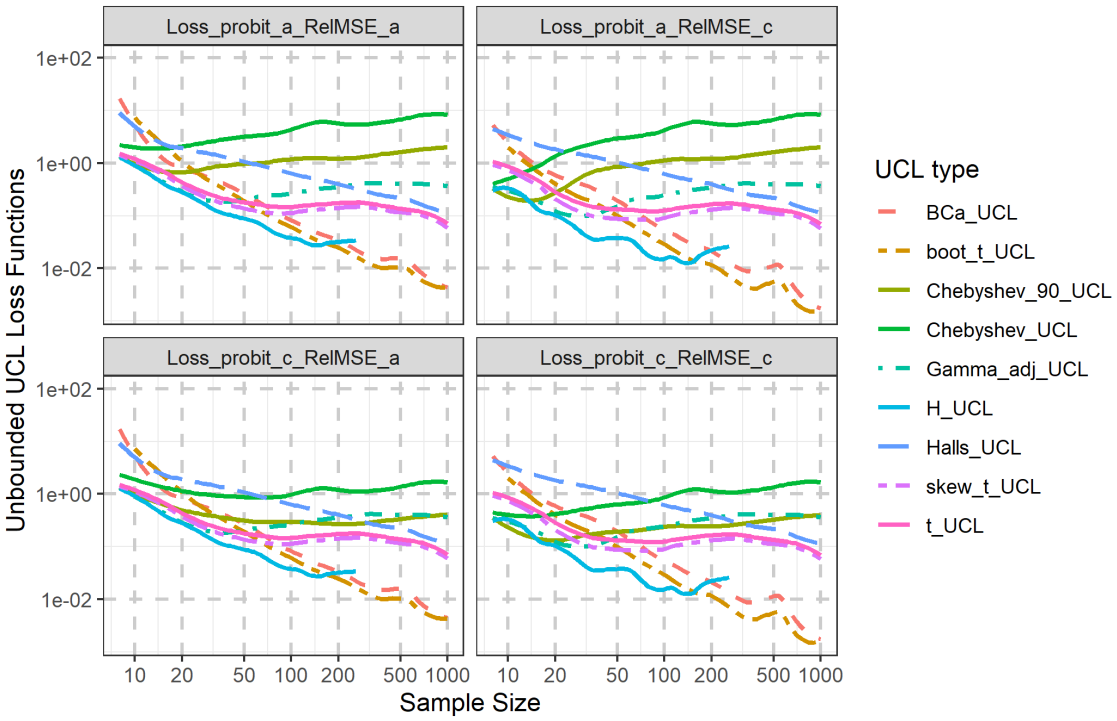

Figure 31. UCL Unbounded Loss with Probit loss for Coverage for Lognormal with Std Dev of Logs in (0.576,0.773]

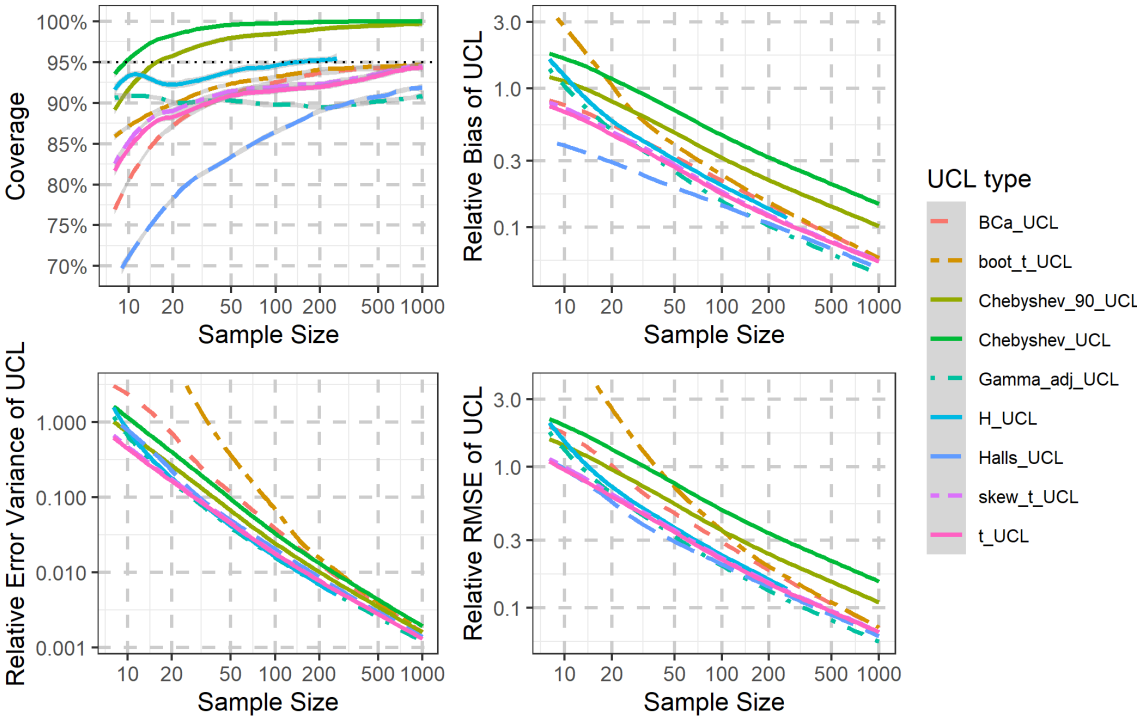

Figure 32. UCL summary for Lognormal with Std Dev of Logs in (0.773,0.982]

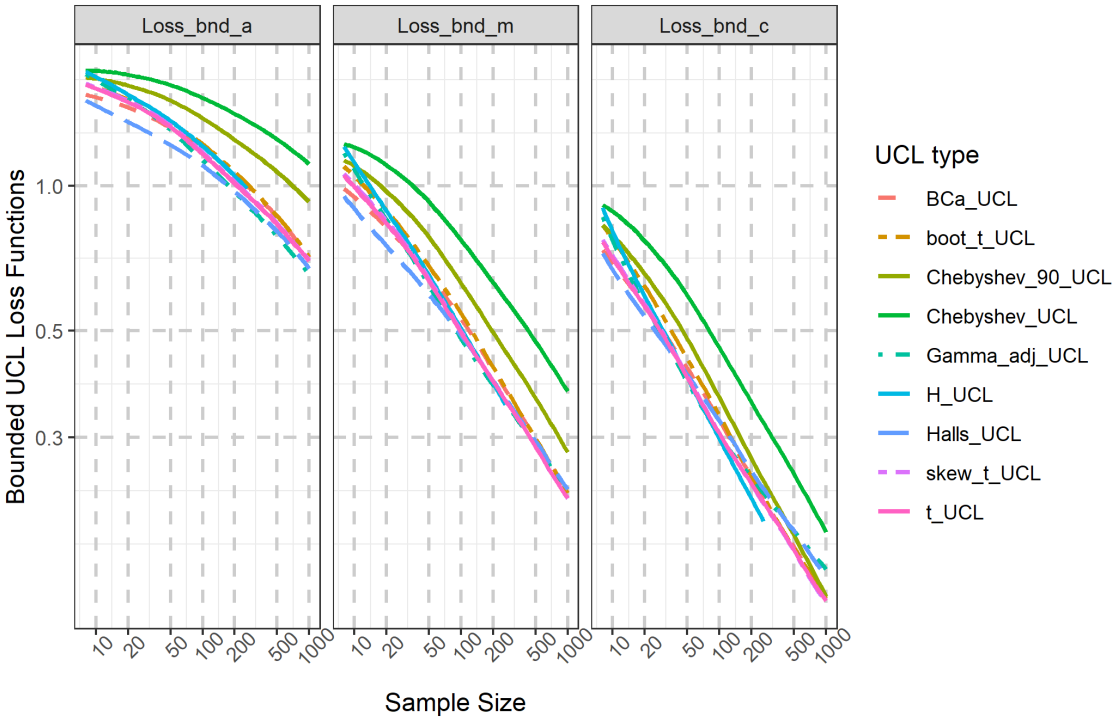

Figure 33. UCL Bounded Loss for Lognormal with Std Dev of Logs in (0.773,0.982]

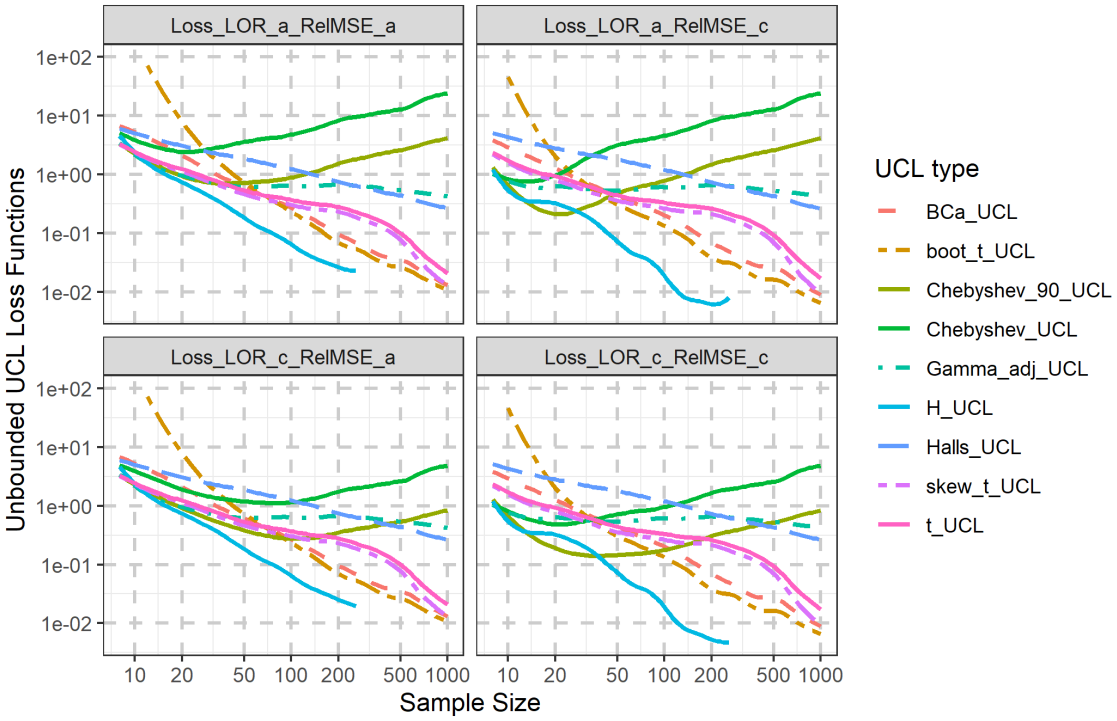

Figure 34. UCL Unbounded Loss with LOR loss for Coverage for Lognormal with Std Dev of Logs in (0.773,0.982]

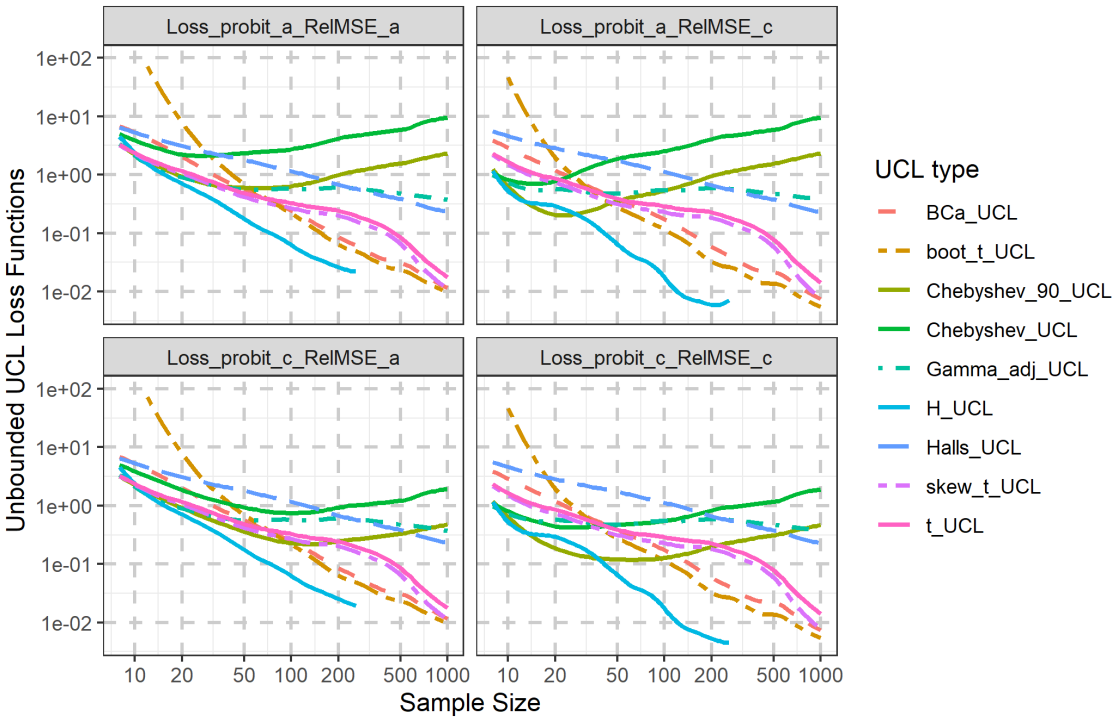

Figure 35. UCL Unbounded Loss with Probit loss for Coverage for Lognormal with Std Dev of Logs in (0.773,0.982]

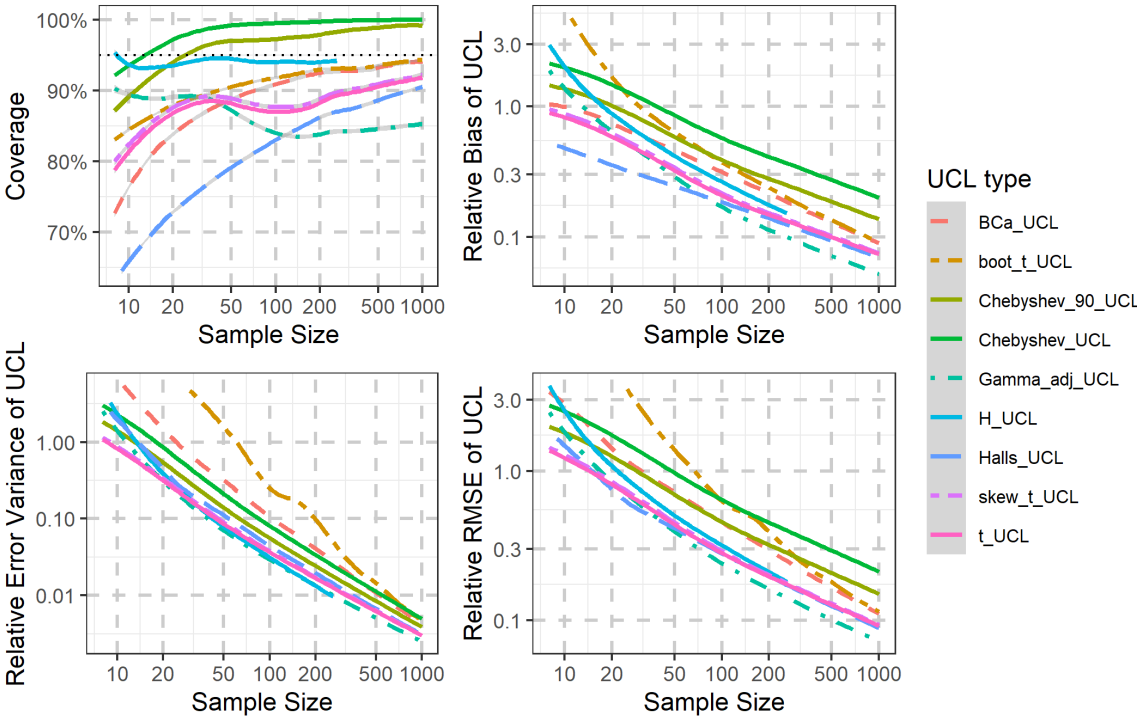

Figure 36. UCL summary for Lognormal with Std Dev of Logs in (0.982,1.17]

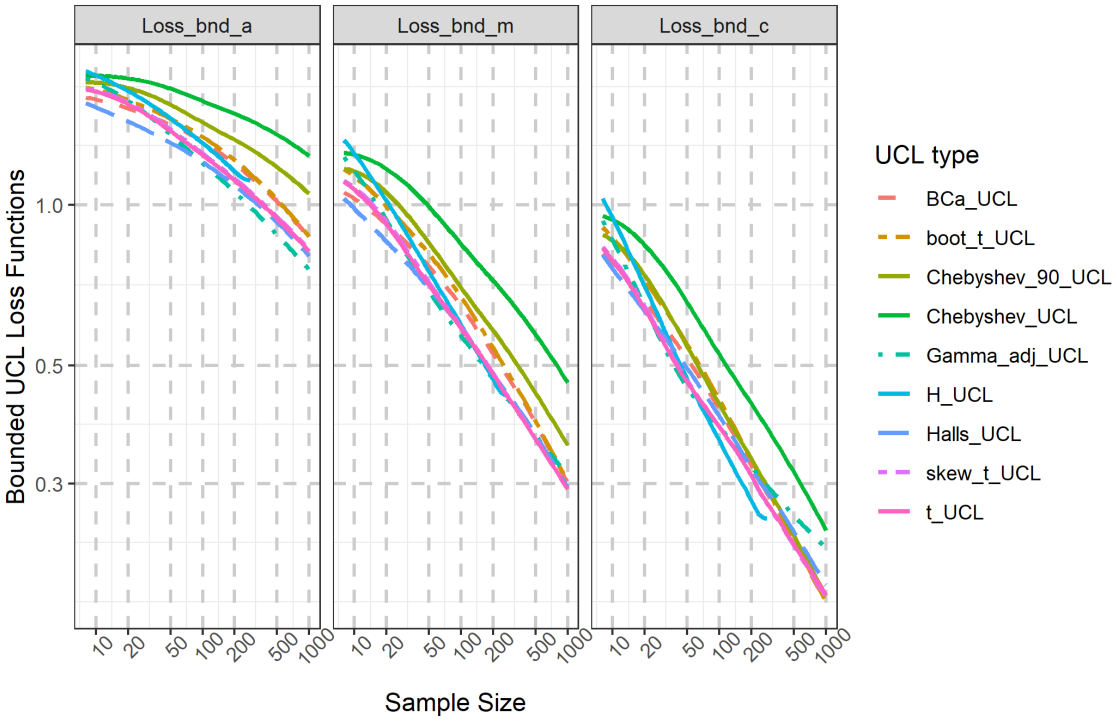

Figure 37. UCL Bounded Loss for Lognormal with Std Dev of Logs in (0.982,1.17]

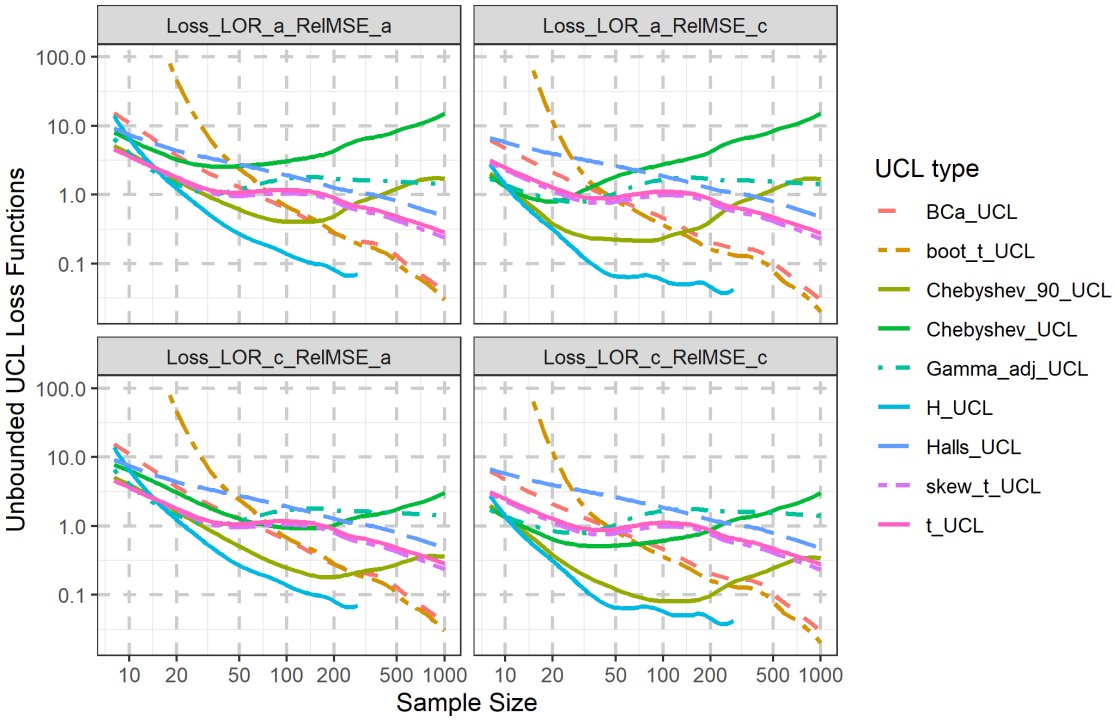

Figure 38. UCL Unbounded Loss with LOR loss for Coverage for Lognormal with Std Dev of Logs in (0.982,1.17]

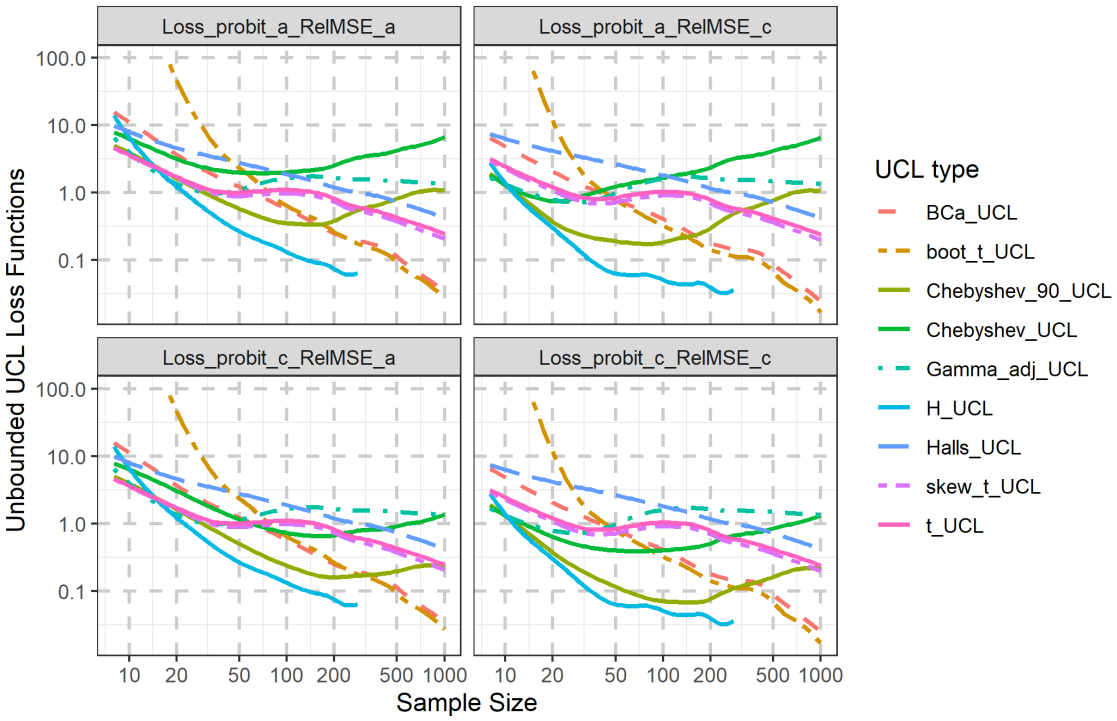

Figure 39. UCL Unbounded Loss with Probit loss for Coverage for Lognormal with Std Dev of Logs in (0.982,1.17]

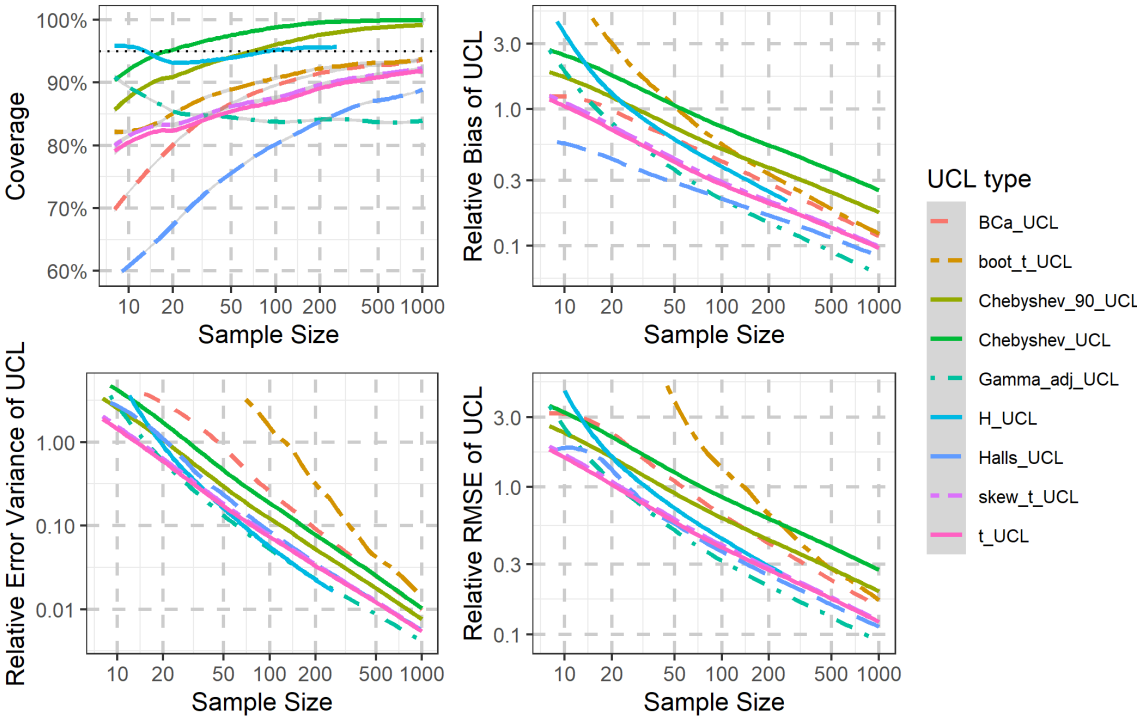

Figure 40. UCL summary for Lognormal with Std Dev of Logs in (1.17,1.37]

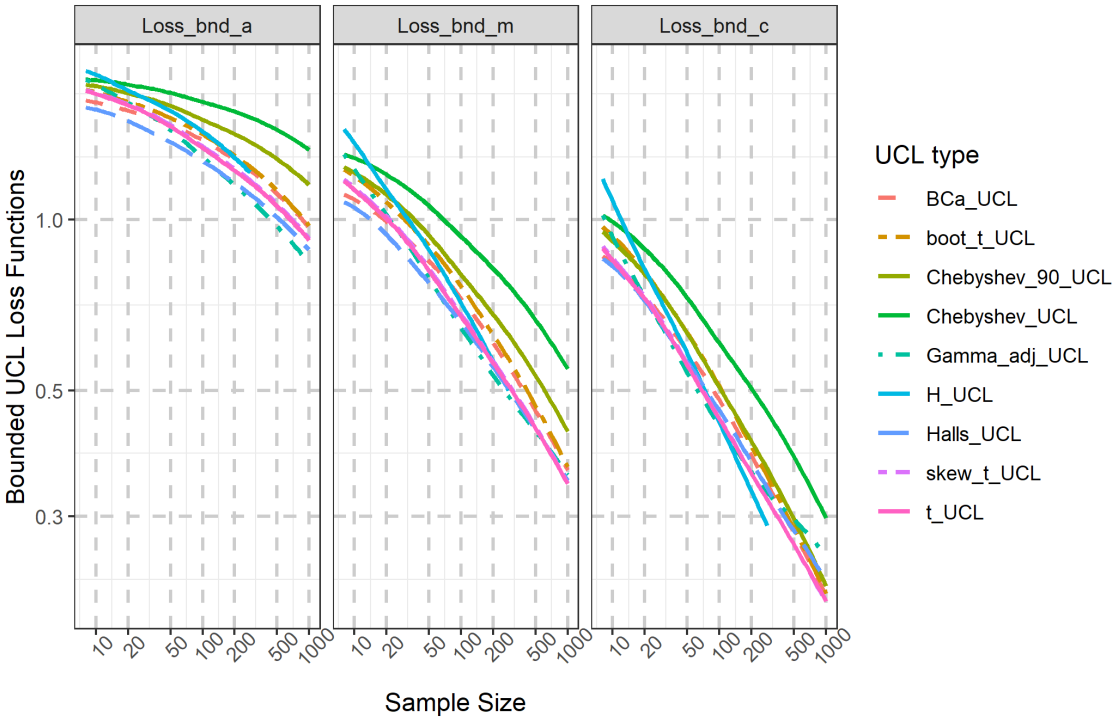

Figure 41. UCL Bounded Loss for Lognormal with Std Dev of Logs in (1.17,1.37]

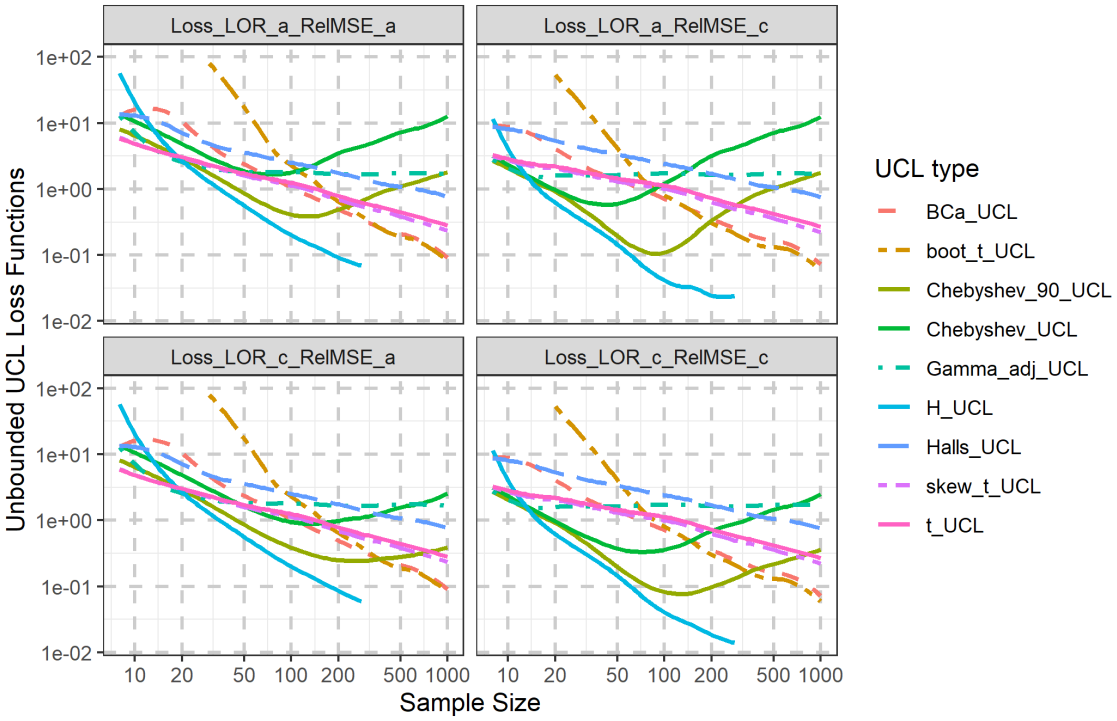

Figure 42. UCL Unbounded Loss with LOR loss for Coverage for Lognormal with Std Dev of Logs in (1.17,1.37]

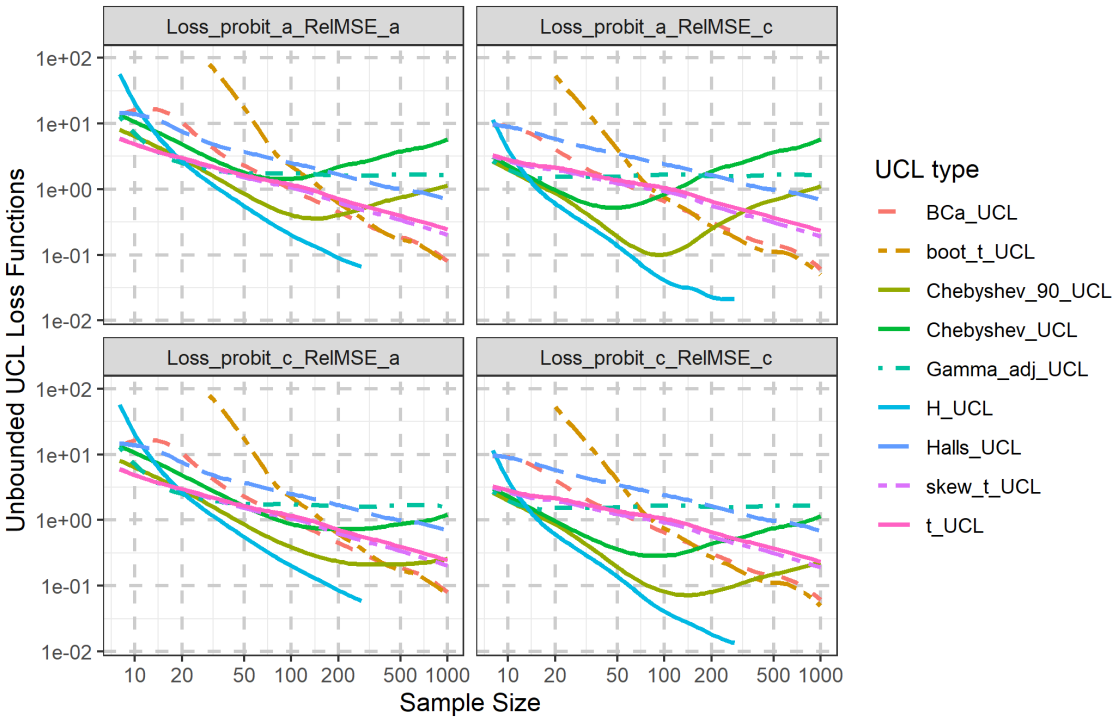

Figure 43. UCL Unbounded Loss with Probit loss for Coverage for Lognormal with Std Dev of Logs in (1.17,1.37]

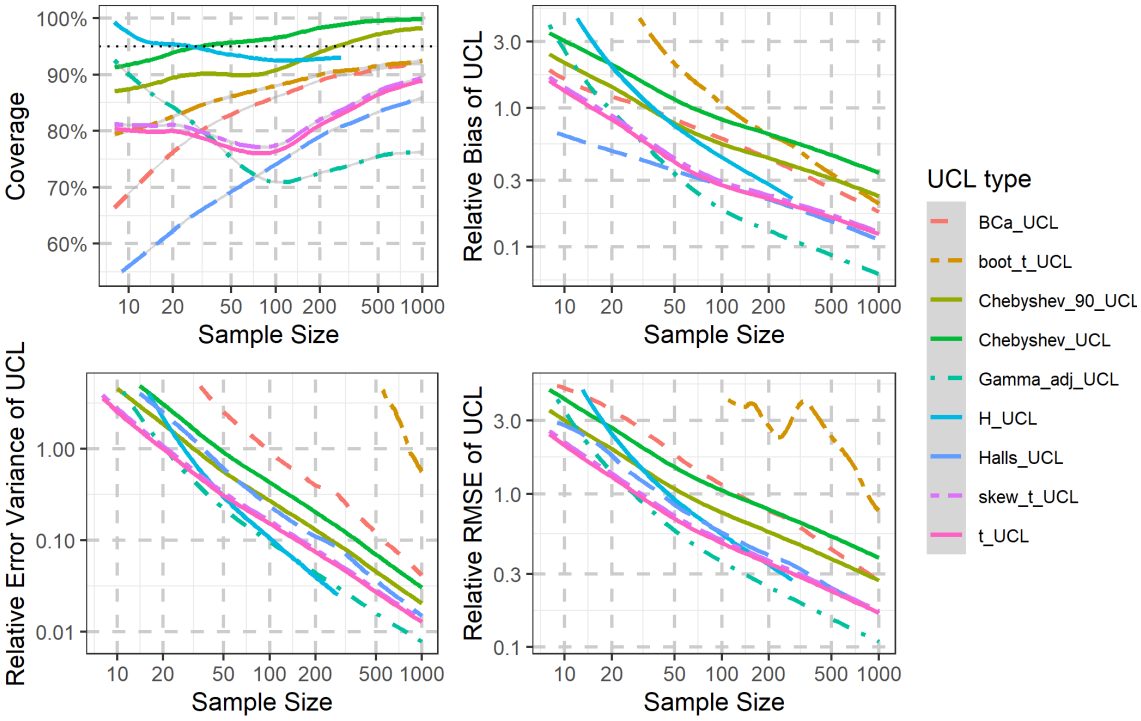

Figure 44. UCL summary for Lognormal with Std Dev of Logs in (1.37,1.57]

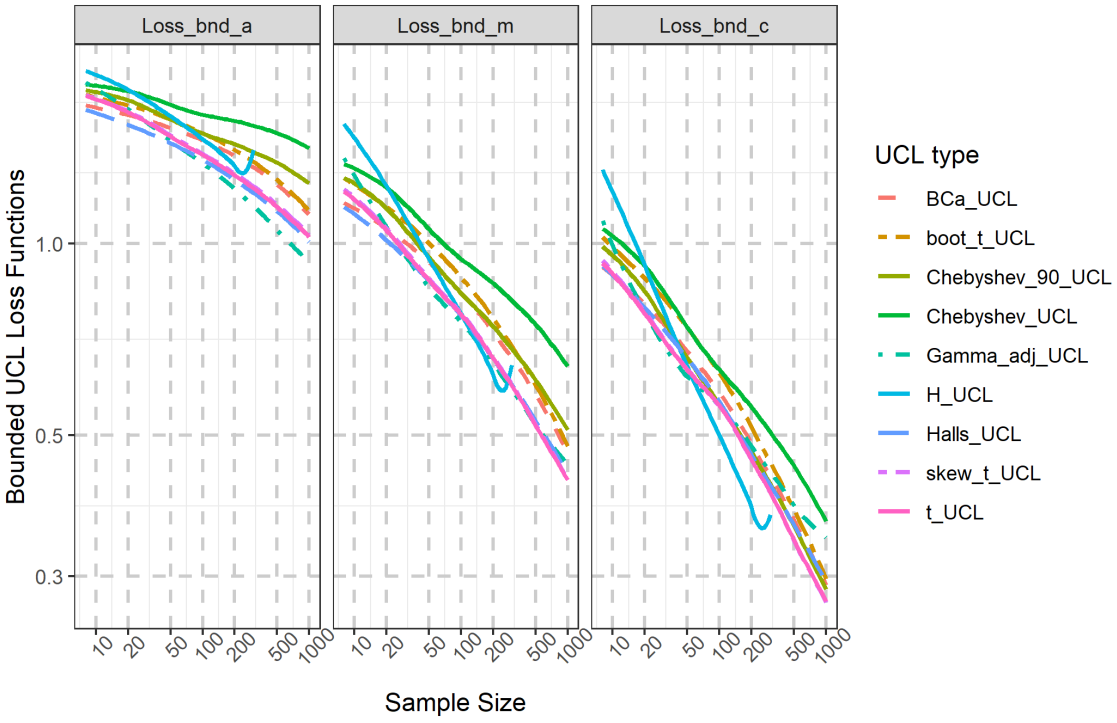

Figure 45. UCL Bounded Loss for Lognormal with Std Dev of Logs in (1.37,1.57]

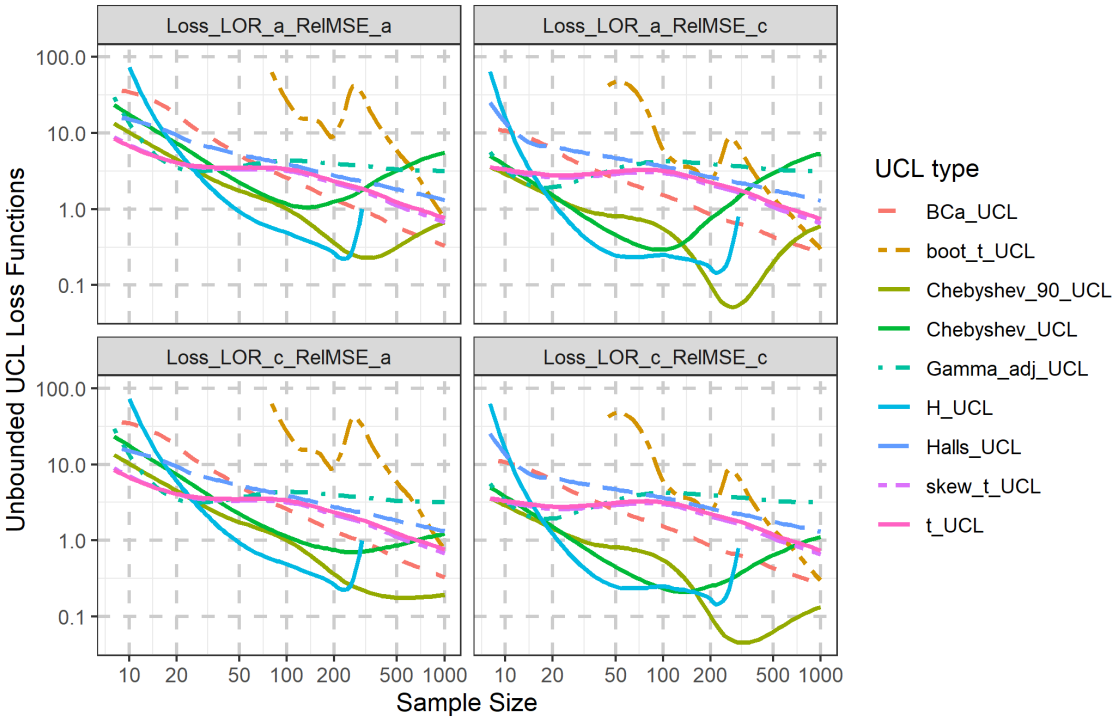

Figure 46. UCL Unbounded Loss with LOR loss for Coverage for Lognormal with Std Dev of Logs in (1.37,1.57]

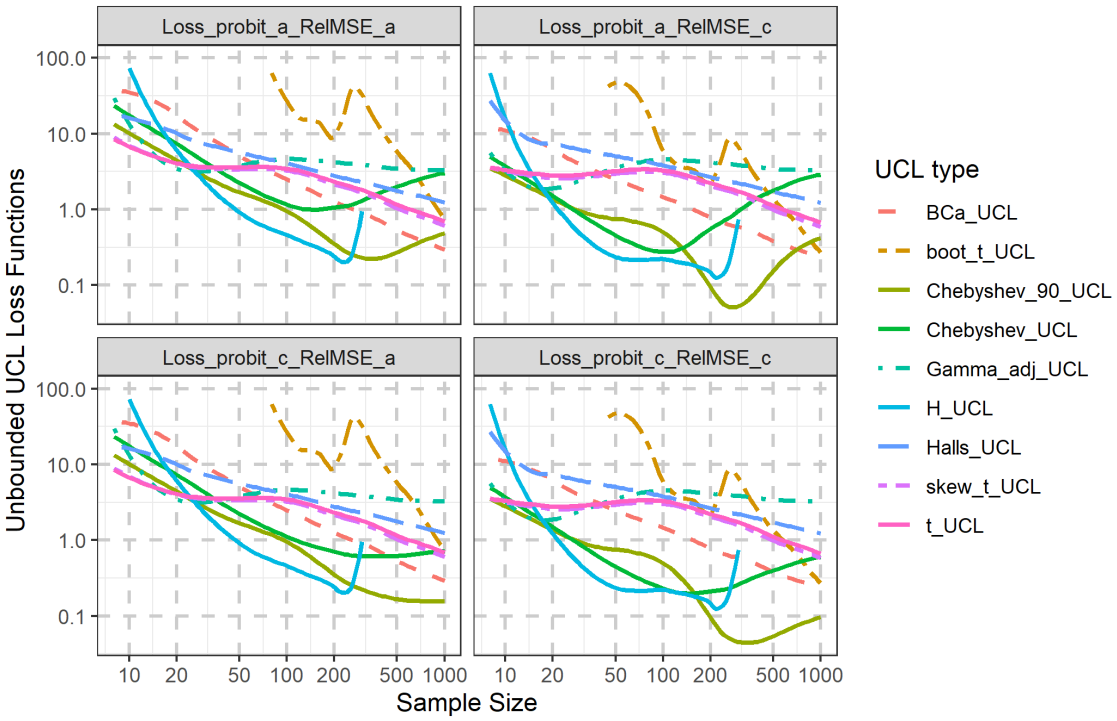

Figure 47. UCL Unbounded Loss with Probit loss for Coverage for Lognormal with Std Dev of Logs in (1.37,1.57]

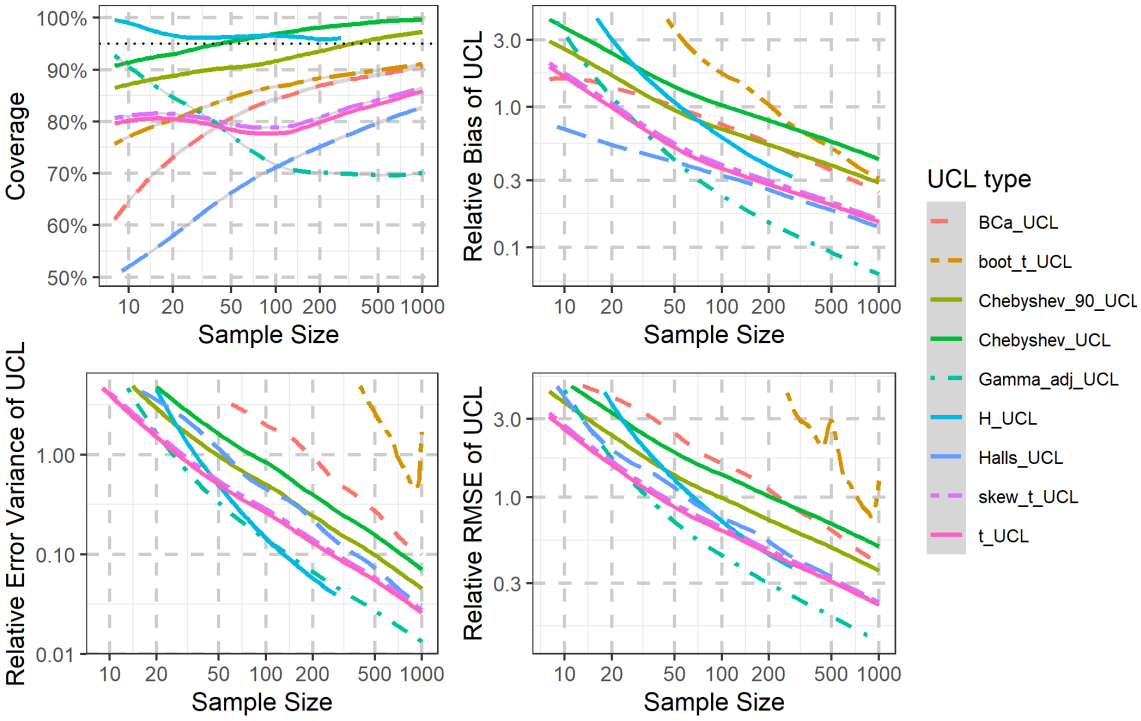

Figure 48. UCL summary for Lognormal with Std Dev of Logs in (1.57,1.73]

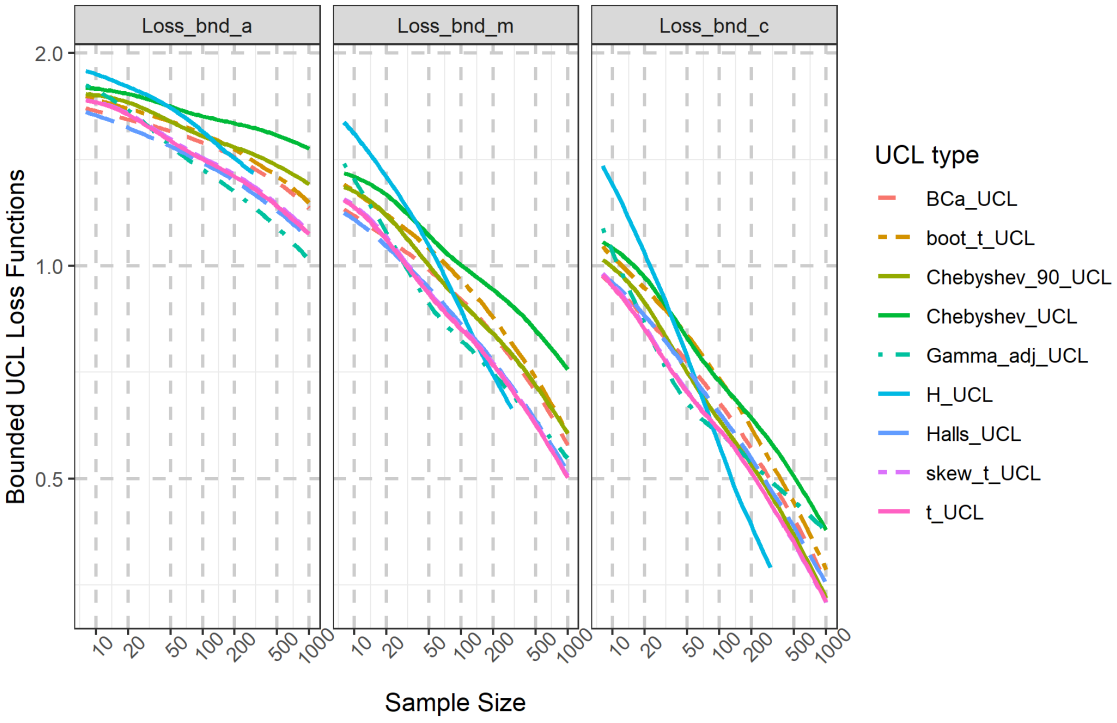

Figure 49. UCL Bounded Loss for Lognormal with Std Dev of Logs in (1.57,1.73]

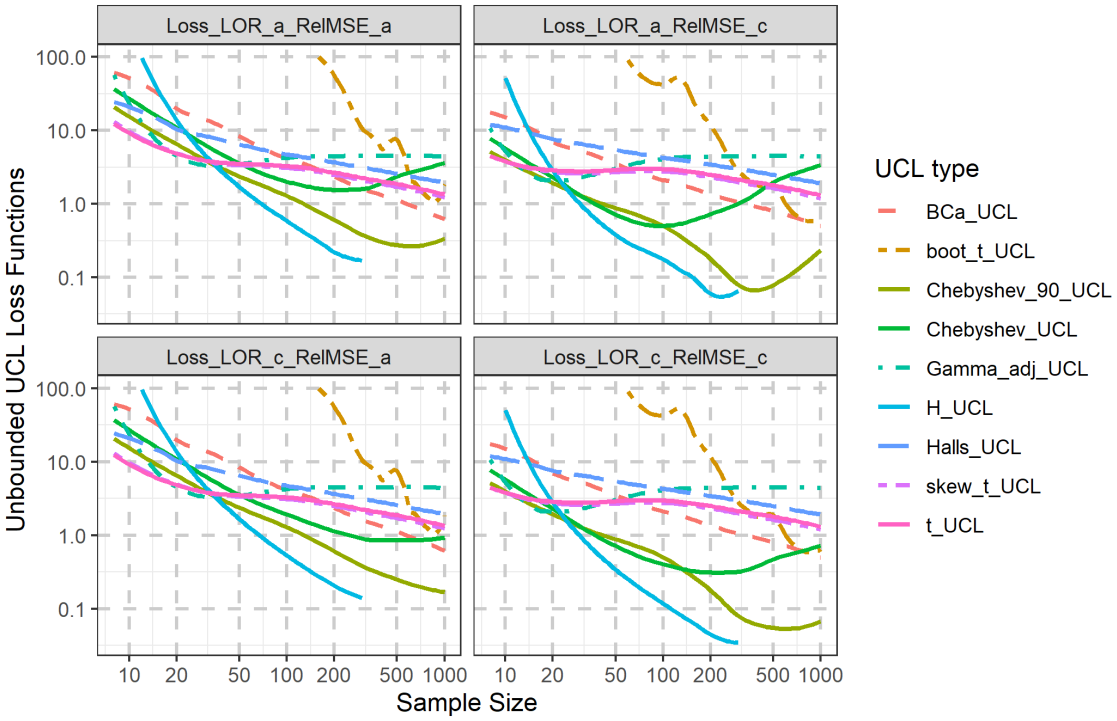

Figure 50. UCL Unbounded Loss with LOR loss for Coverage for Lognormal with Std Dev of Logs in (1.57,1.73]

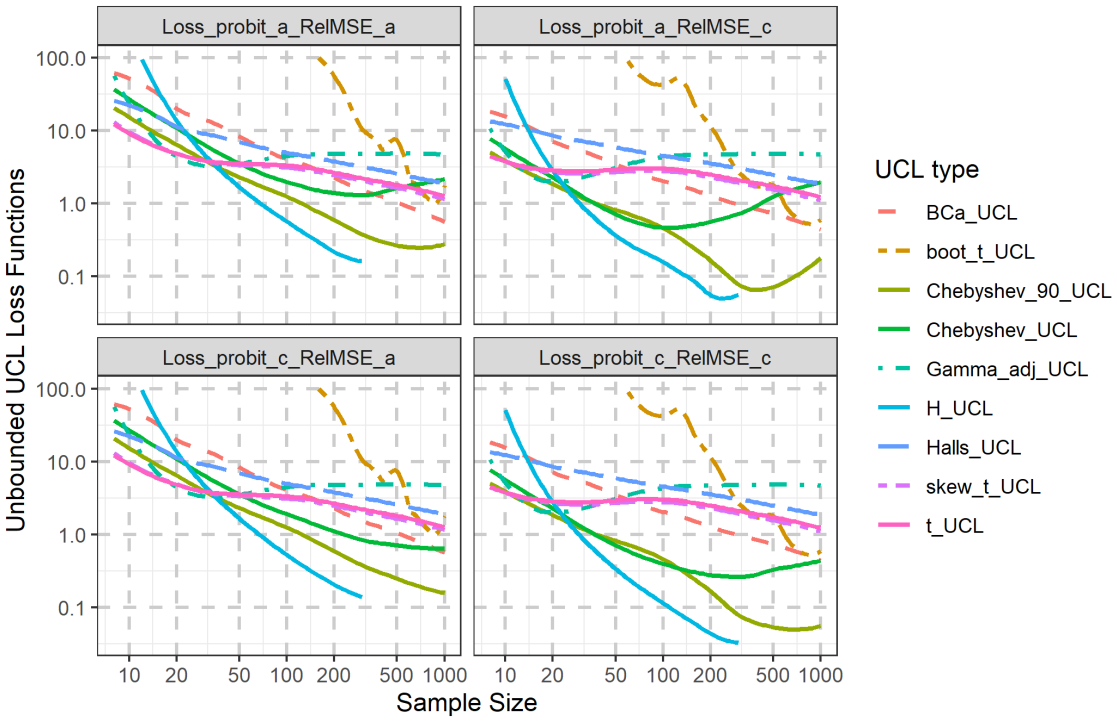

Figure 51. UCL Unbounded Loss with Probit loss for Coverage for Lognormal with Std Dev of Logs in (1.57,1.73]

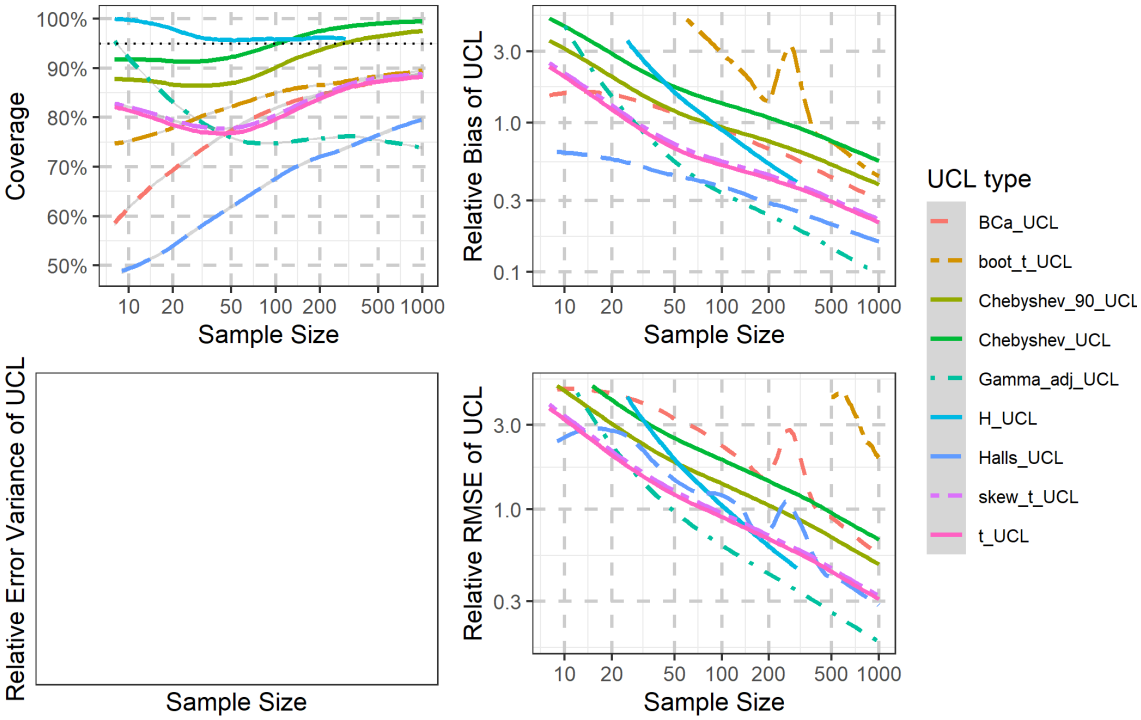

Figure 52. UCL summary for Lognormal with Std Dev of Logs in (1.73,1.95]

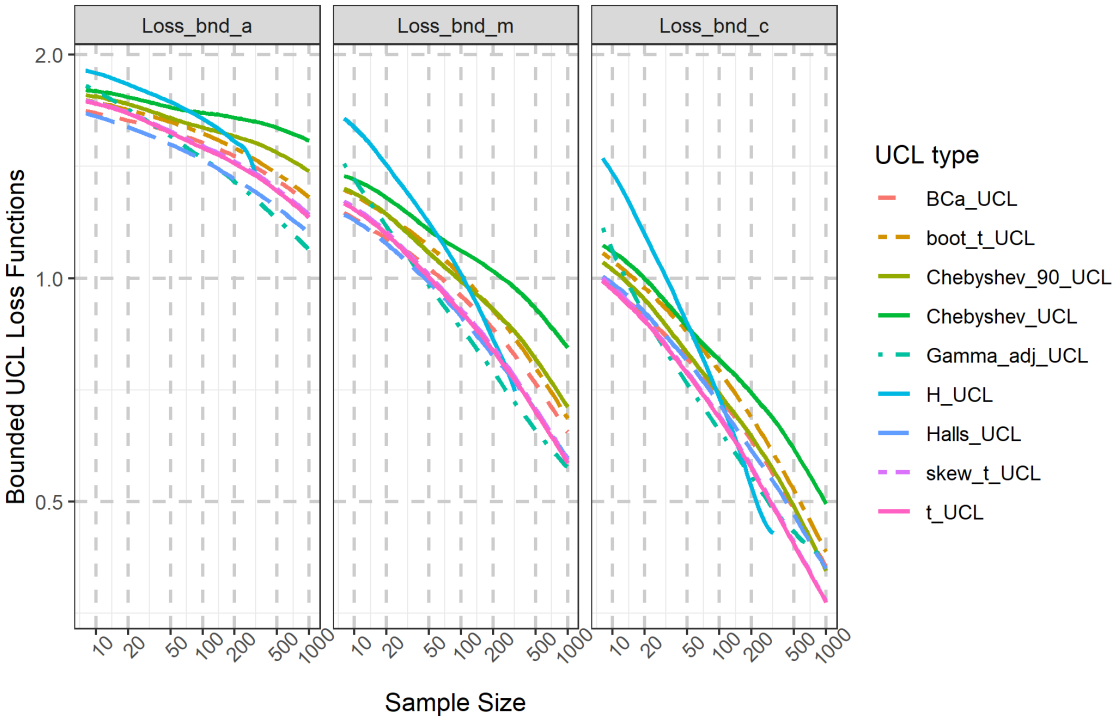

Figure 53. UCL Bounded Loss for Lognormal with Std Dev of Logs in (1.73,1.95]

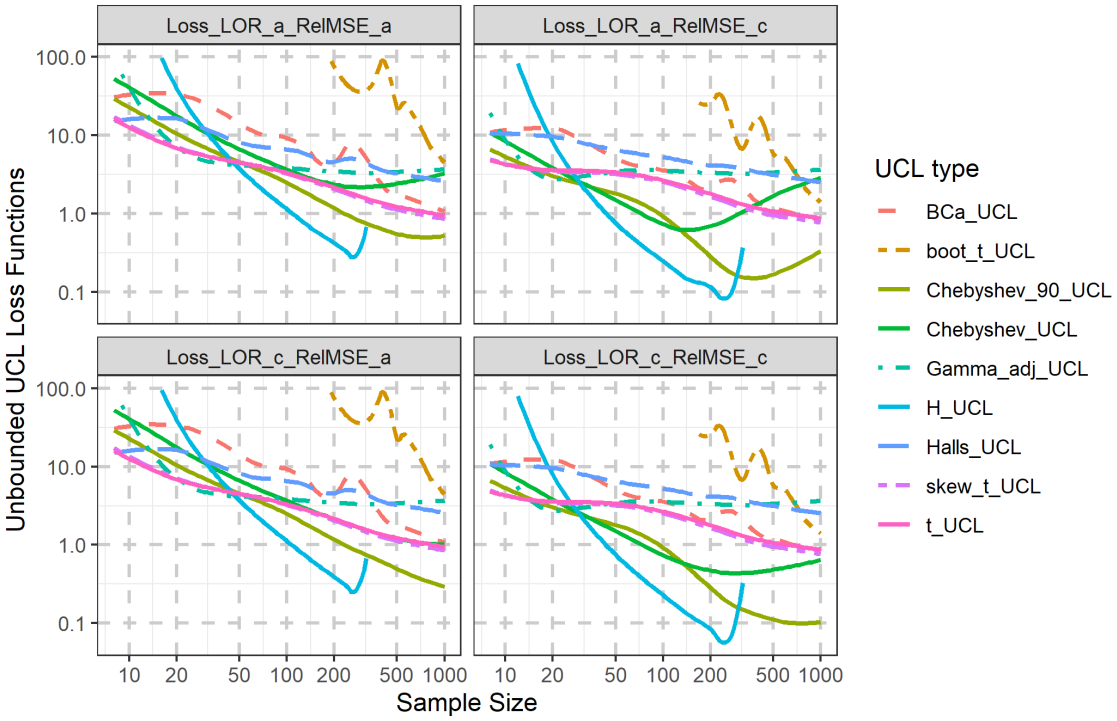

Figure 54. UCL Unbounded Loss with LOR loss for Coverage for Lognormal with Std Dev of Logs in (1.73,1.95]

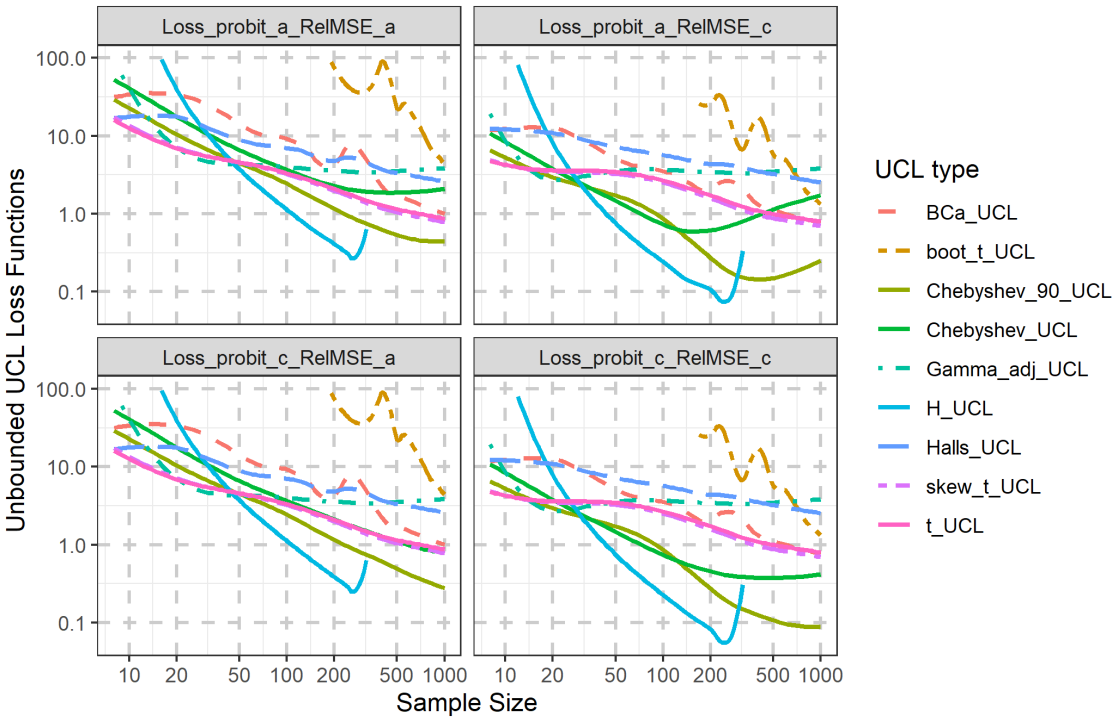

Figure 55. UCL Unbounded Loss with Probit loss for Coverage for Lognormal with Std Dev of Logs in (1.73,1.95]

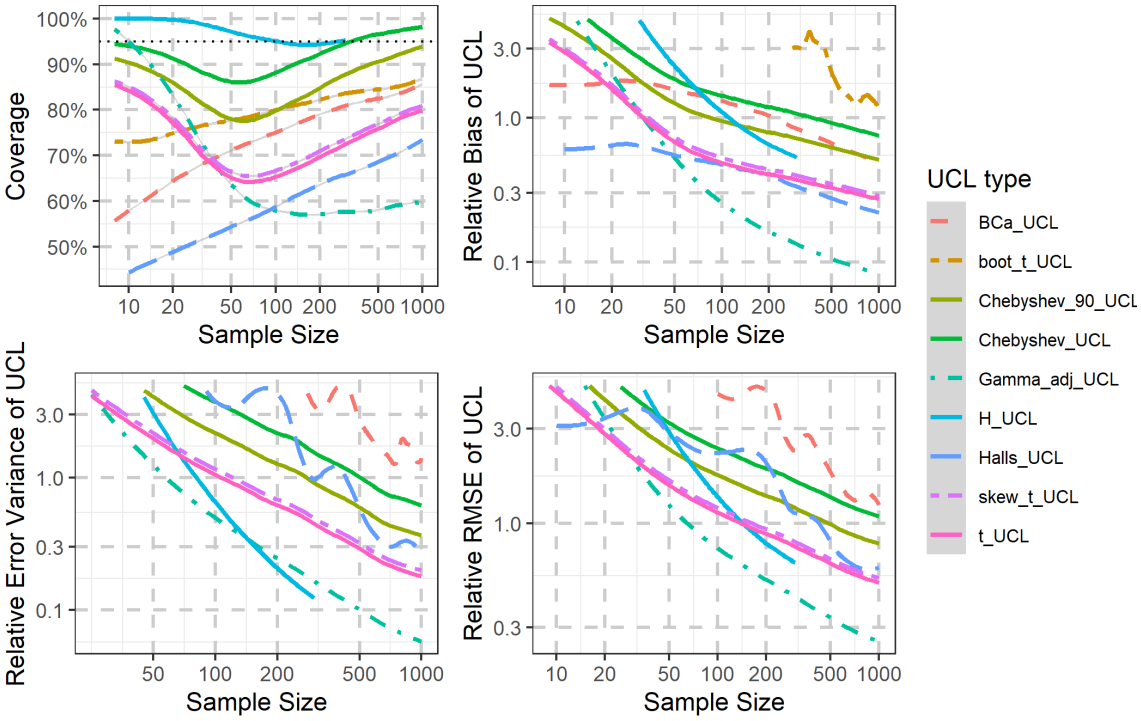

Figure 56. UCL summary for Lognormal with Std Dev of Logs in (1.95,2.25]

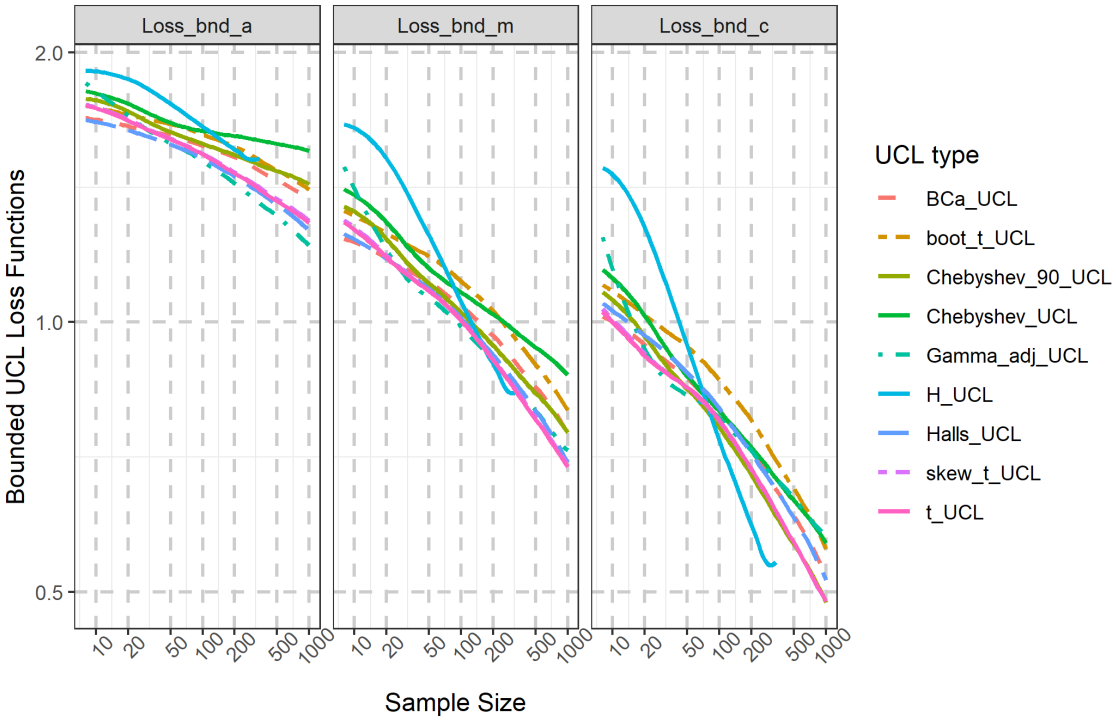

Figure 57. UCL Bounded Loss for Lognormal with Std Dev of Logs in (1.95,2.25]

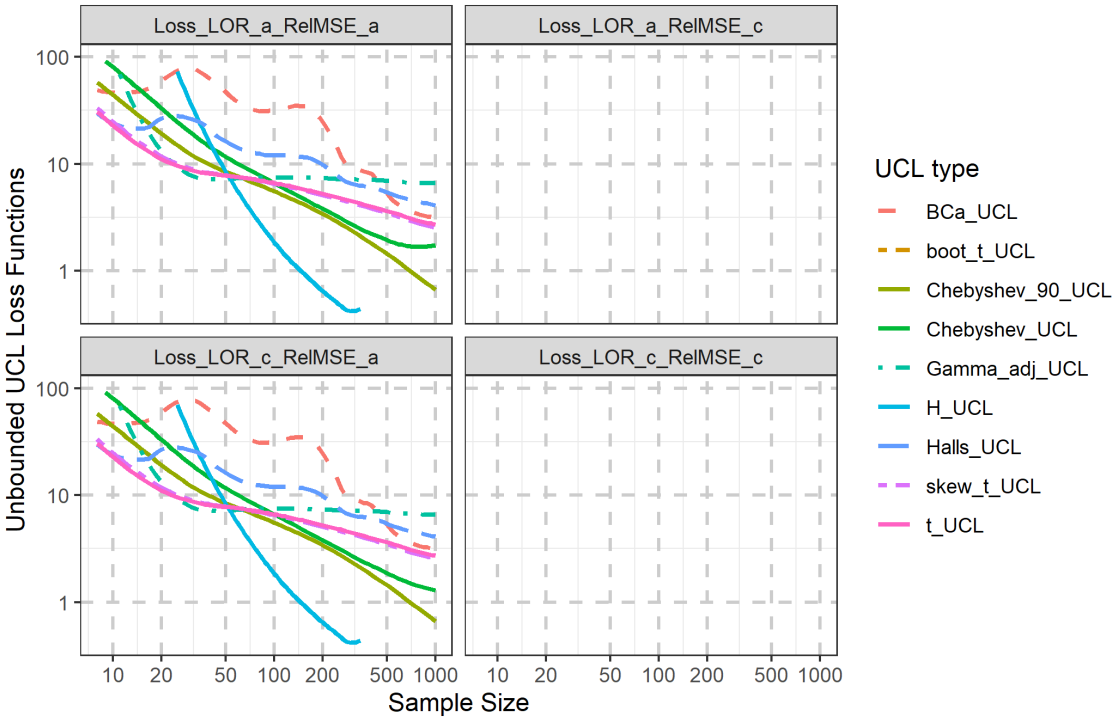

Figure 58. UCL Unbounded Loss with LOR loss for Coverage for Lognormal with Std Dev of Logs in (1.95,2.25]

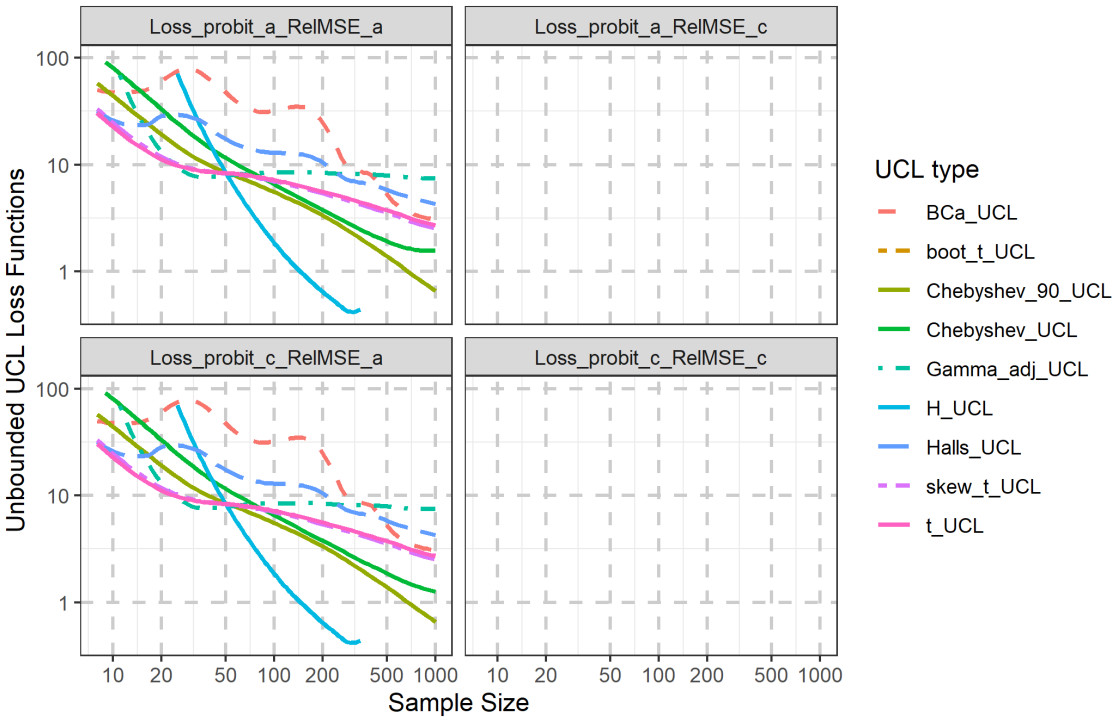

Figure 59. UCL Unbounded Loss with Probit loss for Coverage for Lognormal with Std Dev of Logs in (1.95,2.25]

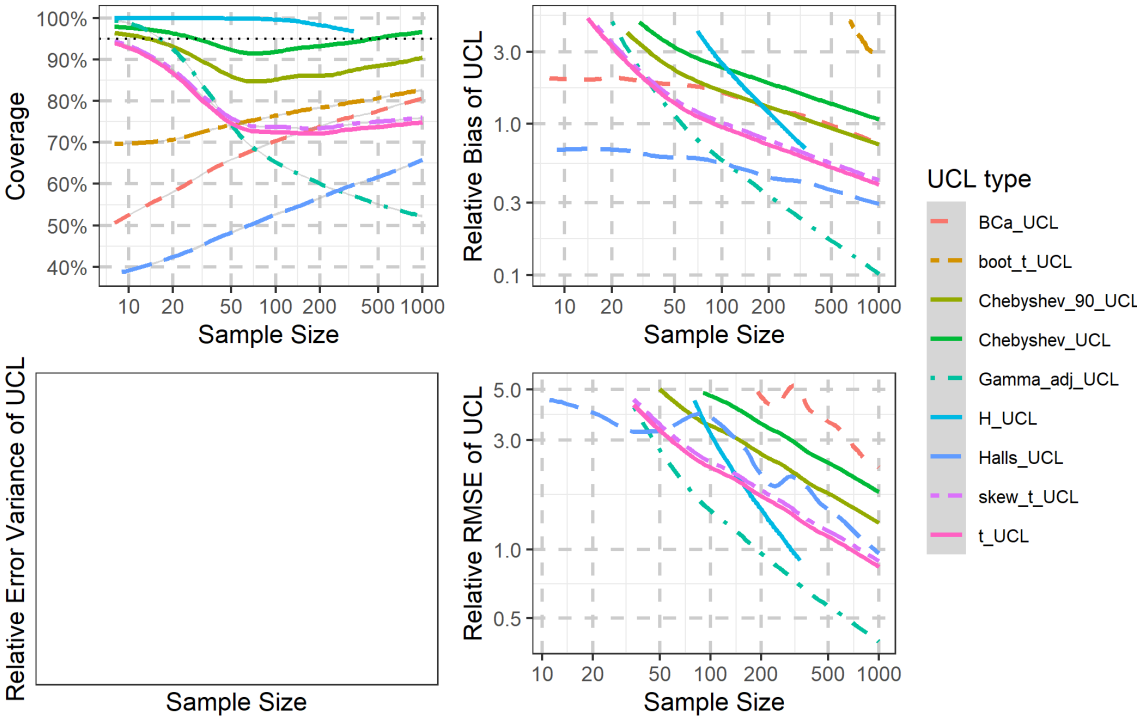

Figure 60. UCL summary for Lognormal with Std Dev of Logs in (2.25,5.41]

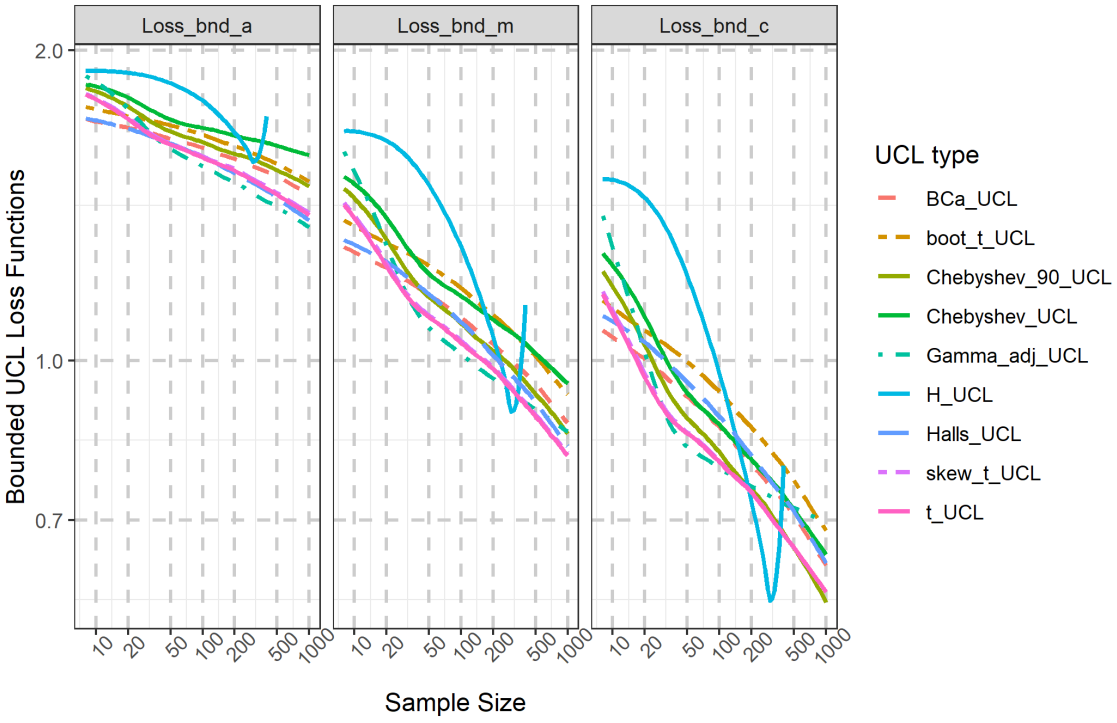

Figure 61. UCL Bounded Loss for Lognormal with Std Dev of Logs in (2.25,5.41]

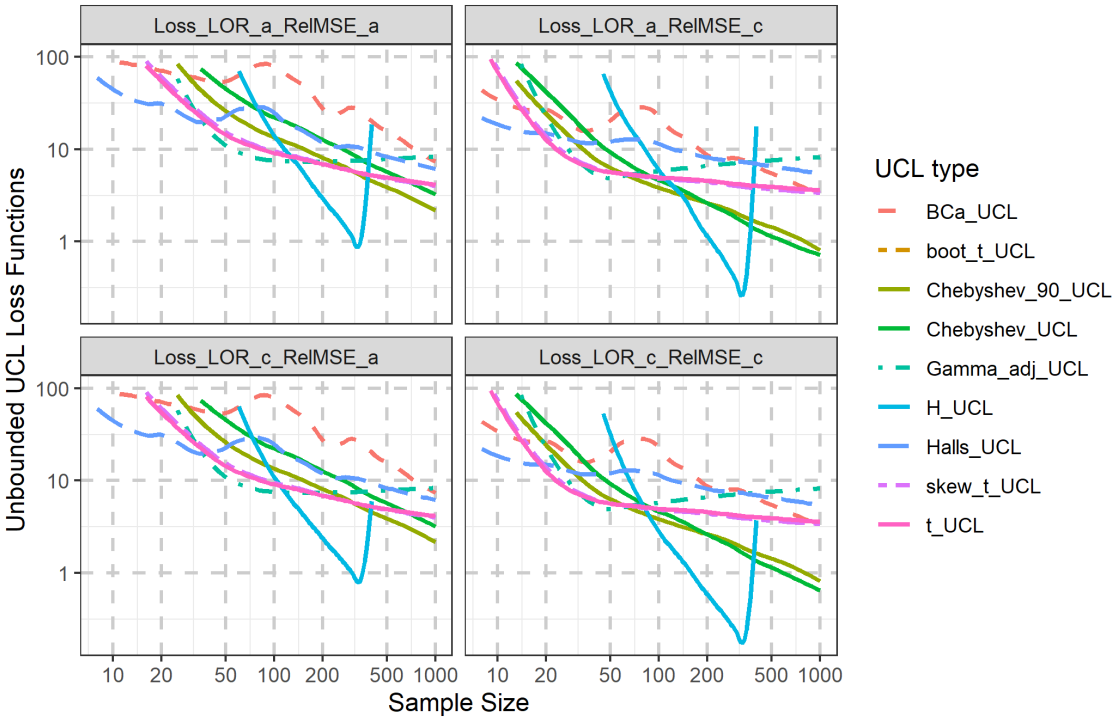

Figure 62. UCL Unbounded Loss with LOR loss for Coverage for Lognormal with Std Dev of Logs in (2.25,5.41]

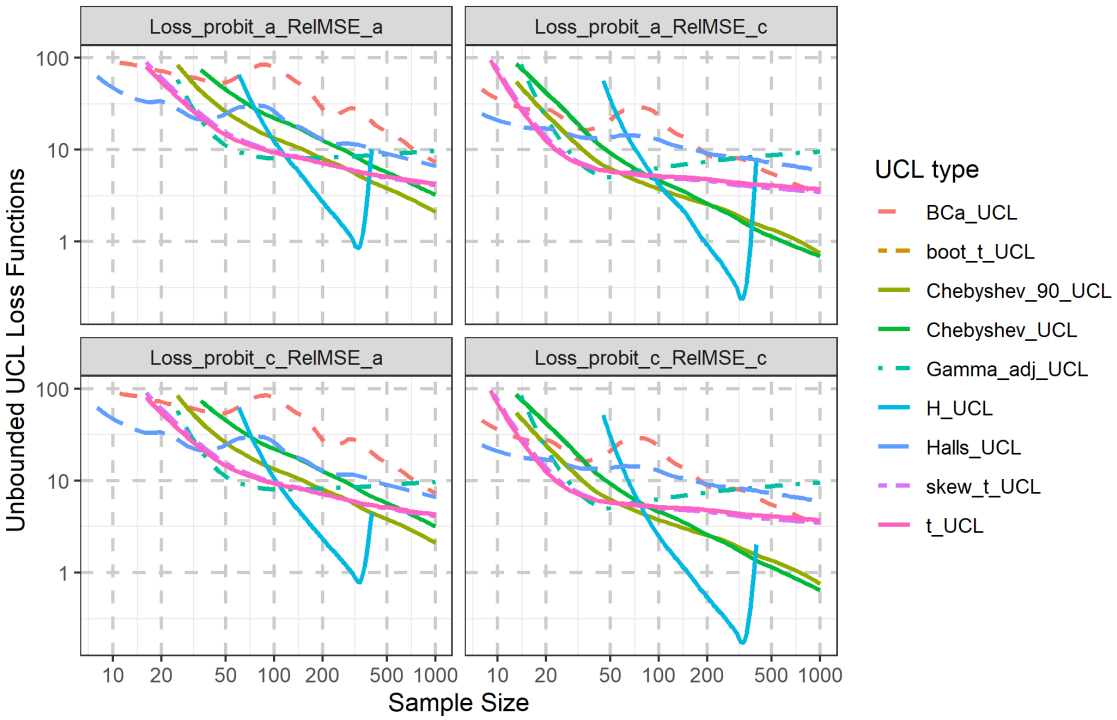

Figure 63. UCL Unbounded Loss with Probit loss for Coverage for Lognormal with Std Dev of Logs in (2.25,5.41]

Appendix B: Session Info

```
## R version 4.1.2 (2021-11-01)
## Platform: x86_64-w64-mingw32/x64 (64-bit)
## Running under: Windows 10 x64 (build 19044)
##
## Matrix products: default
##
## locale:
## [1] LC_COLLATE=English_United States.1252
## [2] LC_CTYPE=English_United States.1252
## [3] LC_MONETARY=English_United States.1252
## [4] LC_NUMERIC=C
## [5] LC_TIME=English_United States.1252
##
## attached base packages:
## [1] stats     graphics  grDevices utils     datasets  methods   base
##
## other attached packages:
##  [1] rpart.plot_3.1.0   rpart_4.1-15       ftExtra_0.2.0      flextable_0.6.10
##  [5] captioner_2.2.3    scales_1.1.1       cowplot_1.1.1      reshape_0.8.8
##  [9] forcats_0.5.1      stringr_1.4.0      dplyr_1.0.7        purrr_0.3.4
## [13] readr_2.1.1        tidyr_1.1.4        tibble_3.1.6       ggplot2_3.3.5
## [17] tidyverse_1.3.1    install.load_1.2.3
##
## loaded via a namespace (and not attached):
##  [1] httr_1.4.2        splines_4.1.2     jsonlite_1.7.2    modelr_0.1.8
##  [5] assertthat_0.2.1  highr_0.9         cellranger_1.1.0  yaml_2.2.1
##  [9] gdtools_0.2.3     lattice_0.20-45   pillar_1.6.4      backports_1.4.1
## [13] glue_1.6.0        uuid_1.0-3        digest_0.6.29     checkmate_2.0.0
## [17] rvest_1.0.2       colorspace_2.0-2  Matrix_1.4-0      htmltools_0.5.2
## [21] plyr_1.8.6        pkgconfig_2.0.3   broom_0.7.10      haven_2.4.3
## [25] officer_0.4.1     tzdb_0.2.0        mgcv_1.8-38       generics_0.1.1
## [29] farver_2.1.0      ellipsis_0.3.2    withr_2.4.3       cli_3.1.0
## [33] magrittr_2.0.1    crayon_1.4.2      readxl_1.3.1      evaluate_0.14
## [37] fs_1.5.2          fansi_0.5.0       nlme_3.1-153      xml2_1.3.3
## [41] tools_4.1.2       data.table_1.14.2 hms_1.1.1         lifecycle_1.0.1
## [45] munsell_0.5.0     reprex_2.0.1      zip_2.2.0         compiler_4.1.2
## [49] systemfonts_1.0.3 rlang_0.4.12      grid_4.1.2        rstudioapi_0.13
## [53] base64enc_0.1-3   labeling_0.4.2    rmarkdown_2.11.3  gtable_0.3.0
## [57] DBI_1.1.2         R6_2.5.1          lubridate_1.8.0   knitr_1.37
## [61] fastmap_1.1.0     utf8_1.2.2        fastmatch_1.1-3   stringi_1.7.6
## [65] Rcpp_1.0.7        vctrs_0.3.8       dbplyr_2.1.1      tidyselect_1.1.1
## [69] xfun_0.29
```

---------- End of Report ----------