A Comparison of Methods for Constructing Population Variance Confidence Intervals

Constructing confidence intervals for population variance using the chi-square distribution is challenging due to its asymmetry, particularly at small sample sizes. A key issue is positioning the interval bounds, as the spacing between the tails differs. This study identifies an optimal method for constructing confidence intervals that balances minimizing interval width and reducing Type II errors. We evaluate three approaches: the equal area tail method, the equal height tail method, and an "optimized" method designed to minimize interval width. The methods are assessed under varying parameters, including sample size (n), significance level (α), true population variance $\left(\sigma^2_{\text{true}}\right)$, and the null hypothesis (H_0). Emphasis is placed on maintaining fixed coverage probabilities and Type I error rates while balancing interval width and Type II errors. Simulation results are presented to compare the methods under different scenarios, highlighting their trade-offs and optimal conditions.

Keywords: Chi-Square Distribution, Confidence Interval, Coverage Probability, Optimization, Significance Level, Type I Error, Type II Error

1 Introduction

Confidence intervals are a fundamental tool for estimating population parameters, such as the true population variance, based on sample data. However, determining the optimal bounds for a confidence interval can be challenging, as there are various methods available for constructing the typical $(1 - \alpha) \times 100$ interval. The asymmetry of the chi-square distribution further complicates the process, making it difficult to identify the most suitable bounds for estimating the population variance.

In this paper, we focus on the chi-square distribution and its application to estimating the true population variance from sample data. The goal is to identify which confidence interval method minimizes the interval width while also reducing Type II error under a fixed significance level, ensuring consistent coverage probability and Type I error. We begin by outlining the general methodology for constructing confidence intervals for population variance, followed by a comparison of the equal area tail, equal height tail, and optimal interval methods. We assess their performance across different parameters, including sample size and significance level. Finally, we conclude with an evaluation of the conditions under which each method proves to be most effective.

Several studies have explored improvements and variations in confidence interval methods. Shorrock (1990) introduced intervals for the population variance that not only depend on the sample variance but also the sample mean, achieving the same length as the shortest interval based solely on the sample variance while providing uniformly higher coverage probability. This extension to the traditional chi-square-based intervals offers more reliable estimates for variance, especially in situations where the sample mean plays a critical role in the variability.

Cojbasic and Loncar (2011) examined one-sided bootstrap-t confidence intervals for variances in skewed distributions, finding that methods such as Hall's transformations provide superior coverage accuracy for lower endpoint intervals, particularly when compared to the standard chi-square-based approaches. These findings emphasize the importance of considering distribution skewness when choosing an interval method. Additionally, Tate and Klett (1959) explored optimal confidence intervals for the variance of a normal distribution, yielding the shortest unbiased intervals for variance estimation. Their work underscores the advantage of using optimized bounds for more efficient estimation in normal distributions.

Building on previous research in the field of confidence intervals, this paper examines different methods for estimating population variance, with a focus on statistical accuracy and practical applicability across various data distributions and sample sizes.

2 Methodology

Below, we begin by outlining the basic construction of the lower and upper bounds for confidence intervals using the chi-square distribution. We then provide a brief overview of key evaluation metrics such as coverage probability, Type I and Type II error rates, and interval width for assessing and comparing the performance of different confidence interval methods. Finally, we detail the calculations involved in constructing the bounds for the equal area tail, equal height tail, and optimal confidence intervals.

2.1 General Case Population Variance Confidence Interval

The chi-square distribution is an important distribution which plays an integral role in statistical inference. It originates when we take the sum of squares of independent and identically distributed random variables. By definition, the chi-square distribution with n degrees of freedom, denoted by χ^2_n, is the distribution of $Z_1^2 + Z_2^2 + \cdots + Z_n^2$ for n independent standard normal variables $Z_1, Z_2, \ldots, Z_n$. In other words,

$$\sum_{i=1}^{n} Z_i^2 \sim \chi^2_n.$$

But before we can formulate a confidence interval for the true variance, σ^2_{true}, we must first prove

$$\sum_{i=1}^{n} \left(Z_i - \bar{Z}\right)^2 \sim \chi^2_{n-1}.$$

We can do so informally by rearranging $\sum_{i=1}^{n} Z_i^2 \sim \chi^2_n$.

$$\begin{aligned}
\sum_{i=1}^{n} Z_i^2 &= \sum_{i=1}^{n} \left(Z_i - \bar{Z} + \bar{Z}\right)^2 \\
&= \sum_{i=1}^{n} \left(\left(Z_i - \bar{Z}\right)^2 + \bar{Z}^2\right) \\
&= \sum_{i=1}^{n} \left(Z_i - \bar{Z}\right)^2 + n \cdot \bar{Z}^2 \\
&= \sum_{i=1}^{n} \left(Z_i - \bar{Z}\right)^2 + \left(\sqrt{n}\right)^2 \cdot \bar{Z}^2 \\
&= \sum_{i=1}^{n} \left(Z_i - \bar{Z}\right)^2 + \left(\sqrt{n} \cdot \bar{Z}\right)^2 \\
&= \sum_{i=1}^{n} \left(Z_i - \bar{Z}\right)^2 + \left(\sqrt{n} \cdot \left(\bar{Z} - 0\right)\right)^2 \\
&= \sum_{i=1}^{n} \left(Z_i - \bar{Z}\right)^2 + \left(\frac{\bar{Z} - 0}{\frac{1}{\sqrt{n}}}\right)^2
\end{aligned}$$

$$\implies \sum_{i=1}^{n} Z_i^2 = \sum_{i=1}^{n} \left(Z_i - \bar{Z}\right)^2 + \left(\frac{\bar{Z} - 0}{\frac{1}{\sqrt{n}}}\right)^2.$$

The left-hand side follows a chi-square distribution with n degrees of freedom, as $\sum_{i=1}^{n} Z_i^2 \sim \chi^2_n$. On the right-hand side, note that $Z_1, Z_2, \ldots, Z_n \overset{i.i.d.}{\sim} Norm\left(0, 1\right)$ which implies that the standard mean $\bar{Z}$ is distributed as $Norm\left(0, \frac{1}{n}\right)$ by the Central Limit Theorem. Standardizing $\bar{Z}$ to a standard normal distribution gives

$$\left(\frac{\bar{Z} - 0}{\frac{1}{\sqrt{n}}}\right) \sim Norm(0, 1).$$

By definition, if $Z_i \sim Norm(0,1)$, then $Z_i^2 \sim \chi_1^2$. Since we just demonstrated $\left(\frac{\bar{Z}-0}{\frac{1}{\sqrt{n}}}\right) \sim Norm(0,1)$, it follows that

$$\left(\frac{\bar{Z}-0}{\frac{1}{\sqrt{n}}}\right)^2 \sim \chi_1^2.$$

Combining our results for the left-hand side and right-hand side, we arrive at

$$\sum_{i=1}^{n} Z_i^2 = \sum_{i=1}^{n}\left(Z_i - \bar{Z}\right)^2 + \left(\frac{\bar{Z}-0}{\frac{1}{\sqrt{n}}}\right)^2$$

$$\implies \chi_n^2 = \sum_{i=1}^{n}\left(Z_i - \bar{Z}\right)^2 + \chi_1^2.$$

To ensure the degrees of freedom are preserved on both sides of the equation, the term $\sum_{i=1}^{n}\left(Z_i - \bar{Z}\right)^2$ must follow a chi-square distribution with $n-1$ degrees of freedom. Therefore

$$\sum_{i=1}^{n}\left(Z_i - \bar{Z}\right)^2 \sim \chi_{n-1}^2. \ \square$$

Using the fact that $\sum_{i=1}^{n}\left(Z_i - \bar{Z}\right)^2 \sim \chi_{n-1}^2$, we construct confidence intervals for the population variance. Let $Y_i \overset{i.i.d.}{\sim} Norm\left(\mu_{\text{true}}, \sigma_{\text{true}}^2\right)$ represent the population distribution, and define $Z_i \overset{i.i.d.}{\sim} Norm(0,1)$ for $i = 1, 2, \ldots, n$. Then, by the Central Limit Theorem, the sample mean satisfies $\bar{Y} \overset{i.i.d.}{\sim} Norm\left(\mu_{\text{true}}, \frac{\sigma_{\text{true}}^2}{n}\right)$, and the mean of the standard normals is $\bar{Z} \sim Norm\left(0, \frac{1}{n}\right)$. This results in

$$\begin{aligned}
\sum_{i=1}^{n}\left(Z_i - \bar{Z}\right)^2 &= \sum_{i=1}^{n}\left(\frac{Y_i - \mu_{\text{true}}}{\sigma_{\text{true}}} - \bar{Z}\right)^2 && \left(Z_i = \tfrac{Y_i - \mu_{\text{true}}}{\sigma_{\text{true}}}\right) \\
&= \sum_{i=1}^{n}\left(\frac{Y_i - \mu_{\text{true}}}{\sigma_{\text{true}}} - \frac{\bar{Y} - \mu_{\text{true}}}{\sigma_{\text{true}}}\right)^2 && \left(\bar{Z} = \tfrac{\bar{Y} - \mu_{\text{true}}}{\sigma_{\text{true}}}\right) \\
&= \frac{1}{\sigma_{\text{true}}^2}\sum_{i=1}^{n}\left(Y_i - \mu_{\text{true}} - \left(\bar{Y} - \mu_{\text{true}}\right)\right)^2 \\
&= \frac{1}{\sigma_{\text{true}}^2}\sum_{i=1}^{n}\left(Y_i - \mu_{\text{true}} - \bar{Y} + \mu_{\text{true}}\right)^2 \\
&= \frac{1}{\sigma_{\text{true}}^2}\sum_{i=1}^{n}\left(Y_i - \bar{Y}\right)^2 \\
&= \frac{(n-1)s^2}{\sigma_{\text{true}}^2} && \left(s^2 = \tfrac{1}{n-1}\textstyle\sum_{i=1}^{n}\left(Y_i - \bar{Y}\right)^2\right) \\
\implies \frac{(n-1)s^2}{\sigma_{\text{true}}^2} &\sim \chi_{n-1}^2.
\end{aligned}$$

To construct a $(1-\alpha)\times 100\%$ confidence interval for σ_{true}^2, we use the fact that a chi-square random variable lies between two critical values with probability $1-\alpha$

$$P\left(\chi_{\text{Lower},\ n-1}^2 \leq \frac{(n-1)s^2}{\sigma_{\text{true}}^2} \leq \chi_{\text{Upper},\ n-1}^2\right) = 1 - \alpha,$$

where $\chi^2_{\text{Lower},\ n-1}$ and $\chi^2_{\text{Upper},\ n-1}$ are the critical values for the lower and upper tails of the chi-square distribution with $n-1$ degrees of freedom. Rearranging for σ^2_{true}, we get

$$
\begin{aligned}
P\left(\chi^2_{\text{Lower},\ n-1} \leq \frac{(n-1)s^2}{\sigma^2_{\text{true}}} \leq \chi^2_{\text{Upper},\ n-1}\right) &= P\left(\frac{1}{\chi^2_{\text{Lower},\ n-1}} \geq \frac{\sigma^2_{\text{true}}}{(n-1)s^2} \geq \frac{1}{\chi^2_{\text{Upper},\ n-1}}\right) \\
&= P\left(\frac{1}{\chi^2_{\text{Upper},\ n-1}} \leq \frac{\sigma^2_{\text{true}}}{(n-1)s^2} \leq \frac{1}{\chi^2_{\text{Lower},\ n-1}}\right) \\
&= P\left(\frac{(n-1)s^2}{\chi^2_{\text{Upper},\ n-1}} \leq \sigma^2_{\text{true}} \leq \frac{(n-1)s^2}{\chi^2_{\text{Lower},\ n-1}}\right)
\end{aligned}
$$

$$
\implies P\left(\frac{(n-1)s^2}{\chi^2_{\text{Upper},\ n-1}} \leq \sigma^2_{\text{true}} \leq \frac{(n-1)s^2}{\chi^2_{\text{Lower},\ n-1}}\right) = 1-\alpha.
$$

Thus, the $(1-\alpha)\times 100\%$ confidence interval for the true population is given by

$$
\left[\frac{(n-1)s^2}{\chi^2_{\text{Upper},\ n-1}}, \frac{(n-1)s^2}{\chi^2_{\text{Lower},\ n-1}}\right].
$$

2.2 Evaluation Metrics

When constructing confidence intervals for population variance, it is essential to assess their performance using well-defined evaluation metrics. These metrics provide insight into how accurately and reliably the confidence intervals estimate the true population variance. Key metrics include coverage probability, which measures the proportion of intervals that capture the true parameter; error rates, which quantify the frequency of incorrect inferences; and interval width, which reflects the precision of the estimate. Together, these criteria offer a comprehensive framework for evaluating the quality and effectiveness of confidence interval methods. Below, we briefly define these metrics in the context of population variance estimation.

Coverage probability is the likelihood that a confidence interval contains the true population variance. For a $(1-\alpha)\times 100\%$ confidence interval, the nominal coverage probability is $(1-\alpha)\times 100\%$. This implies that, on average, $(1-\alpha)\times 100\%$ of such intervals constructed from repeated sampling are expected to include the true population variance. Consequently, for any observed interval, one can be $(1-\alpha)\times 100\%$ confidence that the true population variance lies within its bounds.

A Type I error occurs when the null hypothesis, which is assumed to be true, is incorrectly rejected. The probability of committing a Type I error is denoted by the significance level α. In the context of confidence intervals for population variance, this corresponds to the combined probability of the true variance falling outside the interval bounds, represented by the areas in the tails of the relevant chi-square distribution.

On the other hand, a Type II error arises when the null hypothesis is not rejected despite being false. The probability of a Type II error depends on the specific null hypothesis and the true population variance. Typically, there is an inverse relationship between the probabilities of Type I and Type II errors. The bounds determined by the critical values influence the Type II error rate, which vary depending on the interval construction method.

Interval width measures the precision of a confidence interval and is defined as the difference between its upper and lower bounds. For confidence intervals of population variance, this can be expressed as

$$\frac{(n-1)s^2}{\chi^2_{\text{Lower},\ n-1}} - \frac{(n-1)s^2}{\chi^2_{\text{Upper},\ n-1}},$$

where n is the sample size, s^2 is the sample variance, and $\chi^2_{\text{Upper},\ n-1}$ and $\chi^2_{\text{Upper},\ n-1}$ are the critical values of the chi-square distribution corresponding to the lower and upper bounds, respectively. For a fixed significance level α, the nominal coverage probability and Type I error rate remain consistent across different interval construction methods. However, the interval width varies depending on the choice of critical values, resulting in differences among methods such as the equal area tail, equal height tail, and optimal intervals.

Since interval width and Type II error rates are influenced by the critical values and, by extension, the resulting interval bounds, they serve as our metrics of interest for evaluating and optimizing the performance of confidence interval methods for population variance.

2.3 Equal Area Tail Confidence Interval

As seen in section 2.1, a generic confidence interval for the population variance is defined by the bounds

$$\left[\frac{(n-1)s^2}{\chi^2_{\text{Upper},\ n-1}}, \frac{(n-1)s^2}{\chi^2_{\text{Lower},\ n-1}}\right].$$

In the case of the equal area tail confidence interval, the critical values are chosen such that the area below the lower tail and beyond the upper tail of the chi-square distribution are both equal to $\alpha/2$. Specifically, the lower critical value $\chi^2_{\text{Lower},\ n-1}$ is replaced by $\chi^2_{1-\alpha/2,\ n-1}$, and the upper critical value $\chi^2_{\text{Upper},\ n-1}$ is replaced by $\chi^2_{\alpha/2,\ n-1}$, where the subscripts denote the cumulative area to the **right**. This choice of critical values can be computed easily using statistical software[1] and ensures the condition below holds

$$P\left(\chi^2_{\text{Lower},\ n-1} \le \frac{(n-1)s^2}{\sigma^2_{\text{true}}} \le \chi^2_{\text{Upper},\ n-1}\right) = 1-\alpha.$$

As a result, the equal area tail confidence interval for the population variance can be expressed as

$$\left[\frac{(n-1)s^2}{\chi^2_{\alpha/2,\ n-1}}, \frac{(n-1)s^2}{\chi^2_{1-\alpha/2,\ n-1}}\right].$$

Here, the total area outside the confidence interval is split evenly between the two tails of the chi-square distribution.

2.4 Equal Height Tail Confidence Interval

The equal height tail confidence interval method requires the height of the probability density function (PDF) at the lower critical value to match the height at the upper critical value for the chi-square distribution. Additionally, the sum of the areas in the left and right tails must equal the significance level α. This ensures the confidence interval meets the desired confidence level of $(1-\alpha)\times 100\%$. In other words, to construct the equal height tail confidence interval, two conditions must hold given some lower critical value `x` and some upper critical value `y`:

[1] In `R`, we can use the `qchisq()` function to calculate critical values. The lower critical value $\chi^2_{1-\alpha/2,\ n-1}$ corresponds to the command `qchisq(alpha / 2, n - 1)` while the upper critical value $\chi^2_{\alpha/2,\ n-1}$ corresponds to the command `qchisq(1 - alpha / 2, n - 1)`, where `alpha` denotes the significance level and `n - 1` denotes the degrees of freedom.

1. Equal Height - The chi-square PDF values at the critical points are equal[2] $f(x) = f(y)$.
2. Tail Area - The sum of the tail probabilities equals the significance level $P(\chi^2 < x) + P(\chi^2 > y) = \alpha$.

Using R, this system of equations can be formatted as shown below:

1. `dchisq(x, n - 1)` $=$ `dchisq(y, n - 1)`.
2. `pchisq(x, n - 1)` $+$ `pchisq(y, n - 1, lower.tail = FALSE)` $= \alpha$.

To find the critical values x and y:

- Start with an initial guess, where x is near 0 and y is a large value with equivalent height. This should satisfy the equal height condition but not the tail area condition.
- Then, adjust x and y iteratively until the tail area condition is satisfied. This involves increasing x to grow the left tail area $P(\chi^2 < x)$ while decreasing y to increase the right tail area $P(\chi^2 > y)$ and maintain the equal height condition.

Initially, with x near 0 and y large, the sum of the tail areas is approximately 0. However, as x and y approach the median of the distribution, the sum of the left and right tail areas monotonically increases toward 1. By the Intermediate Value Theorem, since α is a value between 0 and 1, there must exist some critical value pair x and y such that the tail areas sum to α. The final critical points that satisfy both conditions are what we want.

Alternatively, we could simply use a root-solver. Regardless, our solution will yield the unique critical value pair $\chi^2_{\text{HLower},n-1}$ and $\chi^2_{\text{HUpper},n-1}$. Plugging these into the confidence interval bounds, the equal height tail confidence interval is

$$\left[\frac{(n-1)s^2}{\chi^2_{\text{HUpper},\ n-1}}, \frac{(n-1)s^2}{\chi^2_{\text{HLower},\ n-1}}\right].$$

2.5 Optimal Confidence Interval

To derive the optimal $(1-\alpha) \times 100\%$ confidence interval for the population variance, the goal is to minimize the expected width of the interval. This involves solving an optimization problem to find the lower and upper critical values of the chi-square distribution that minimize the interval width. In other words,

$$\underset{\chi^2_{\text{Lower},\ n-1},\ \chi^2_{\text{Upper},\ n-1}}{\text{argmin}} \quad E[w],$$

where w, width, is defined as the difference between the lower and upper bounds of the interval. It is given by

$$\frac{(n-1)s^2}{\chi^2_{\text{Lower},\ n-1}} - \frac{(n-1)s^2}{\chi^2_{\text{Upper},\ n-1}}.$$

[2] Note how the the chi-square distribution is strictly decreasing for degrees of freedom less than 3, resulting in the equal height condition not holding. It may not hold for large significance levels either.

The expected width, $E[w]$, can then be expressed as

$$\begin{aligned}
E[w] &= E\left[\frac{(n-1)s^2}{\chi^2_{\text{Lower},\ n-1}} - \frac{(n-1)s^2}{\chi^2_{\text{Upper},\ n-1}}\right] \\
&= E\left[(n-1)\,s^2\right] \cdot \left(\frac{1}{\chi^2_{\text{Lower},\ n-1}} - \frac{1}{\chi^2_{\text{Upper},\ n-1}}\right) \\
&= (n-1) \cdot E\left[s^2\right] \cdot \left(\frac{1}{\chi^2_{\text{Lower},\ n-1}} - \frac{1}{\chi^2_{\text{Upper},\ n-1}}\right) \\
&= (n-1) \cdot \sigma^2_{\text{true}} \cdot \left(\frac{1}{\chi^2_{\text{Lower},\ n-1}} - \frac{1}{\chi^2_{\text{Upper},\ n-1}}\right)
\end{aligned}$$

where $n-1$ is the degrees of freedom and σ^2_{true} is the true population variance. The critical values $\chi^2_{\text{Lower},\ n-1}$ and $\chi^2_{\text{Upper},\ n-1}$ are determined by the quantiles q_1' and q_2' on the chi-square distribution such that $0 \leq q_1' < 0.5 < q_2' \leq 1$, with the constraint $q_1 + (1 - q_2) = \alpha$[3]. These quantiles ensure that the total area in the two tails of the chi-square distribution equals α.

Using this formulation of the expected width, the minimization problem simplifies to

$$\underset{\chi^2_{\text{Lower},\ n-1},\ \chi^2_{\text{Upper},\ n-1}}{\operatorname{argmin}} \quad (n-1) \cdot \sigma^2_{\text{true}} \cdot \left(\frac{1}{\chi^2_{\text{Lower},\ n-1}} - \frac{1}{\chi^2_{\text{Upper},\ n-1}}\right).$$

Note how $(n-1)$ and σ^2_{true} are constants - the optimization problem further reduces to

$$\underset{\chi^2_{\text{Lower},\ n-1},\ \chi^2_{\text{Upper},\ n-1}}{\operatorname{argmin}} \left(\frac{1}{\chi^2_{\text{Lower},\ n-1}} - \frac{1}{\chi^2_{\text{Upper},\ n-1}}\right).$$

This difference must remain positive due to the ordering of the chi-square quantiles

$$\begin{aligned}
&\chi^2_{\text{Lower},\ n-1} < \chi^2_{\text{Upper},\ n-1} \\
\implies &\frac{1}{\chi^2_{\text{Upper},\ n-1}} < \frac{1}{\chi^2_{\text{Lower},\ n-1}} \\
\implies &0 < \frac{1}{\chi^2_{\text{Lower},\ n-1}} - \frac{1}{\chi^2_{\text{Upper},\ n-1}}.
\end{aligned}$$

To minimize the expected width $E(w)$, the term

$$\left(\frac{1}{\chi^2_{\text{Lower},\ n-1}} - \frac{1}{\chi^2_{\text{Upper},\ n-1}}\right)$$

should be as close to zero as possible while satisfying the constraints of the confidence interval.

By increasing the q_1' (closer to 0.5), the critical value $\chi^2_{\text{Lower},\ n-1}$ necessarily increases (since it is determined by the quantile), which reduces the term $\frac{1}{\chi^2_{\text{Lower},\ n-1}}$. Conversely, decreasing q_2' (closer to 0.5) lowers $\chi^2_{\text{Upper},\ n-1}$, thereby increasing the term $\frac{1}{\chi^2_{\text{Upper},\ n-1}}$. When both q_1' and q_2' approach 0.5, these terms converge, and their difference approaches zero

$$\lim_{q_1'^- \to 0.5} \frac{1}{\chi^2_{\text{Lower},\ n-1}} = \lim_{q_2'^+ \to 0.5} \frac{1}{\chi^2_{\text{Upper},\ n-1}}$$

[3]In R, $\chi^2_{\text{Lower},\ n-1} =$ `qchisq(q1, n - 1)` and $\chi^2_{\text{Upper},\ n-1} =$ `qchisq(q2, n - 1)` for $q_1' =$`q1` and $q2' =$`q2`.

$$\implies \lim_{q_1'^{-} \to 0.5} \frac{1}{\chi^2_{\text{Lower},\ n-1}} - \lim_{q_2'^{+} \to 0.5} \frac{1}{\chi^2_{\text{Upper},\ n-1}} = 0.$$

However, the constraint $q_1' + (1 - q_2') = \alpha$ must be satisfied to ensure a $(1-\alpha) \cdot 100\%$ confidence interval. This requirement means that increasing q_1' forces an increase in q_2' by the same amount, and vice versa, making it impossible to independently optimize one quantile without affecting the other.

Unlike the symmetric normal distribution, where equal-tailed confidence intervals have quantiles symmetrically placed at $\alpha/2$ and $1-\alpha/2$, the chi-square distribution's right skewness complicates the process. The optimal quantiles cannot be determined using equal-area tails and must instead be found by solving the optimization problem numerically.

Optimization algorithms are well-suited for this task, efficiently identifying the values of q_1' and q_2' that correspond to the critical values $\chi^2_{\text{Lower},\ n-1}$ and $\chi^2_{\text{Upper},\ n-1}$, minimizing $E[w]$ while satisfying the constraints $0 \leq q_1' < 0.5 < q_2' \leq 1$ and $q_1 + (1 - q_2) = \alpha$.

Denoting the optimal critical values obtained from this optimization as $\chi^2_{\text{OLower},\ n-1}$ and $\chi^2_{\text{OUpper},\ n-1}$, the resulting confidence interval is given by

$$\left[\frac{(n-1)s^2}{\chi^2_{\text{OUpper},\ n-1}}, \frac{(n-1)s^2}{\chi^2_{\text{OLower},\ n-1}}\right].$$

3 Results

In the following subsections, we explore the impact of varying parameters such as sample size and significance level on the performance of our three confidence interval methods of interest. Simulations of confidence intervals are conducted and graphically compared across metrics to determine which methods are ideal for each scenario.

3.1 Sample Size

For each type of confidence interval method — equal area tail, equal height tail, and optimal — we conduct 10000 replications, where each replication involves the construction of a confidence interval for a randomly generated sample. Do this while fixing sample sizes at 10, then at 25, then at 50, and finally at 100. The samples are generated with a true population variance of 1 and a significance level of 0.05.

From this, we assess the performance of each confidence interval method across the 10000 replications for each available sample size. The evaluation metrics include the average width of the confidence intervals (Mean Width), the variability between the widths (SD Width), the average proportion of intervals that correctly captured the true population variance (Mean Coverage), and the variability between the coverage probabilities (SD Coverage).

The tables provided summarize these metrics for each confidence interval method / sample size combination. Each row represents the results of 10000 replications for a given sample size, where the confidence intervals are constructed via the specified method for each replication at the provided sample size.

Table 1: Equal Area Confidence Interval Where ($m = 10000, \sigma^2_{\text{true}} = 1, \alpha = 0.05$)

Sample Size	Mean Width	SD Width	Mean Coverage	SD Coverage
10	2.8627	0.013507	0.9482	0.002216
25	1.3330	0.003866	0.9470	0.002240
50	0.8545	0.001706	0.9518	0.002142
100	0.5777	0.000817	0.9493	0.002194

Table 2: Equal Height Confidence Interval Where ($m = 10000, \sigma^2_{\text{true}} = 1, \alpha = 0.05$)

Sample Size	Mean Width	SD Width	Mean Coverage	SD Coverage
10	4.2173	0.019899	0.9491	0.002198
25	1.4860	0.004310	0.9472	0.002236
50	0.8972	0.001792	0.9517	0.002144
100	0.5913	0.000836	0.9502	0.002175

Table 3: Optimal Confidence Interval Where ($m = 10000, \sigma^2_{\text{true}} = 1, \alpha = 0.05$)

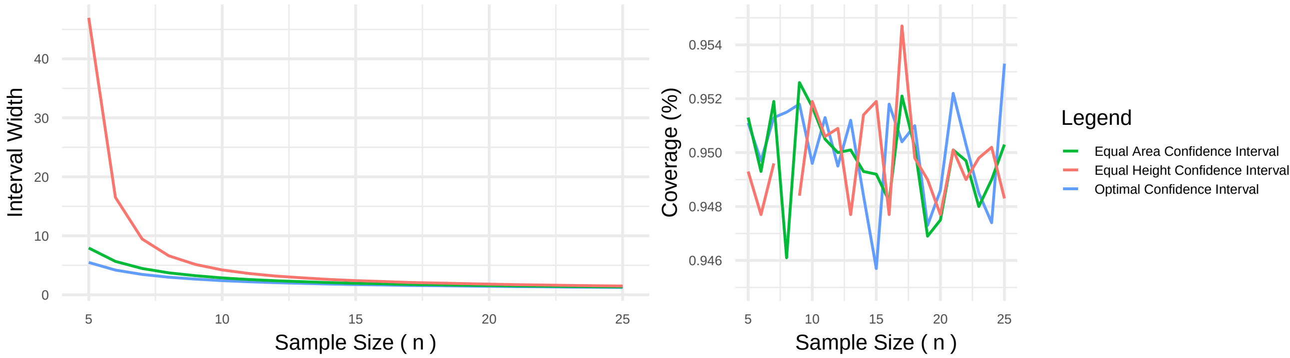
Sample Size	Mean Width	SD Width	Mean Coverage	SD Coverage
10	2.3983	0.011316	0.9481	0.002218
25	1.2427	0.003604	0.9517	0.002144
50	0.8289	0.001655	0.9516	0.002146
100	0.5736	0.000811	0.9496	0.002188

Across all confidence interval methods, we observe that both the Mean Width and SD Width decrease as sample size increases. This behavior aligns with expectations, as the chi-square distribution becomes increasingly symmetric with larger sample sizes. As a result, the differences between the lower and upper bounds of the intervals diminish, and the bounds across methods converge. This convergence is consistent with the behavior of the normal distribution, where the equal area tail, equal height tail, and optimal confidence intervals become equivalent due to its inherent symmetry. Meanwhile, the coverage probability remains stable around 95% (regardless of sample size or method) since it is entirely determined by our fixed significance level of $\alpha = 0.05$.

When comparing the individual rows across the tables, it is clear that the average width of the confidence intervals for the optimal method (Table 3) is notably smaller than that of the other methods, especially at smaller sample sizes. Furthermore, the average width of the confidence intervals for the equal area tail method (Table 1) is consistently smaller than that of the equal height tail method. However, as the sample size increases, the differences in interval width between the methods become much smaller.

To further investigate patterns, we plot sample size against average interval width and sample size against average coverage probability across a broad range of sample sizes ranging from 5 to 100.

In Figure 1**(a)**, the equal height tail confidence interval displays significantly larger widths at small sample sizes but gradually converges with the other methods as sample size increases. The equal area tail confidence interval is sandwiched between the two other methods while the optimal confidence interval consistently achieves the tightest interval across sample sizes.



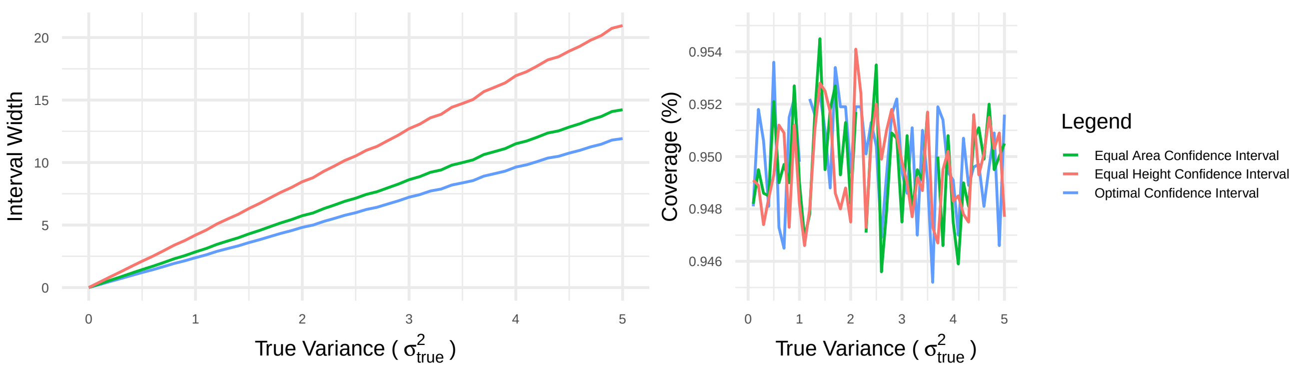
(a) Confidence Interval Widths as a Function of n **(b)** Coverage of Confidence Intervals Across n Values

Figure 1: Performance of Confidence Intervals for Varying Sample Size (n) Values for $\left(m = 10000, \sigma^2_{\text{true}} = 1, \alpha = 0.05\right)$

In Figure 1**(b)**, the average coverage probability remains stable at around 95% for all methods, regardless of sample size or methodology.

3.2 True Variance

To evaluate the performance of the different confidence interval methods across varying true variance values, we perform 10000 replications for each method / true variance combination, where the methods are equal area tail, equal height tail, and optimal while the true variance ranges from 0 to 5 (inclusive and by 0.1). In each replication, the sample size remains constant at 10, with the significance level fixed at 0.05. The plot shown below illustrates the effects of true variance on interval width and coverage over this range of true variances.



(a) Confidence Interval Widths as a Function of σ^2_{true} **(b)** Coverage of Confidence Intervals Across σ^2_{true} Values

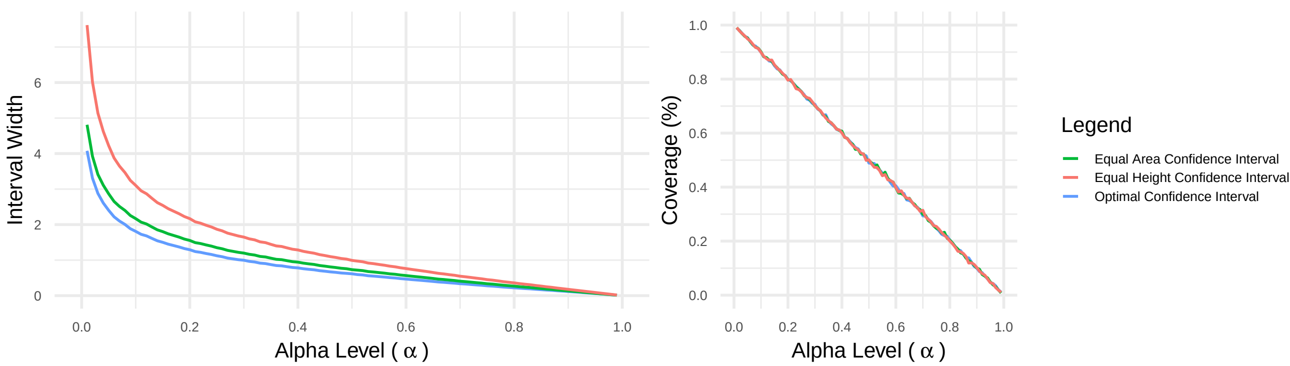
Figure 2: Performance of Confidence Intervals for Varying True Variance $\left(\sigma^2_{\text{true}}\right)$ Values for $(m = 10000, n = 10, \alpha = 0.05)$

In Figure 2**(a)**, the equal height tail confidence interval displays significantly larger widths at large true variance values but gradually converges with the other methods as true variance decreases. The equal area tail confidence interval is sandwiched between the two other methods while the optimal confidence interval consistently achieves the tightest interval across true variances. For all methods, the relationship between interval width and true variance appears to be linear.

In Figure 2(**b**), the average coverage probability remains stable at around 95% for all methods, regardless of true population variance or methodology.

3.3 Significance Level

To assess the performance of the confidence interval methods across a range of significance levels, we conduct 10000 replications for each method — equal area tail, equal height tail, and optimal — at significance levels varying from 0 to 1. For each replication, a confidence interval is constructed using a random sample with the given significance level. The sample size remains constant at 10 across all replications, and the true variance is held at 1. Plotting this data, we can assess the impact of varying significance level on the interval width and coverage probability.



(a) Confidence Interval Widths as a Function of α **(b)** Coverage of Confidence Intervals Across α Values

Figure 3: Performance of Confidence Intervals for Varying Alpha Level (α) Values for $\left(m = 10000, n = 10, \sigma^2_{\text{true}} = 1\right)$

In Figure 3(**a**), the equal height tail confidence interval displays significantly larger widths at small significance levels but gradually converges with the other methods as significance level increases. The equal area tail confidence interval is sandwiched between the two other methods while the optimal confidence interval consistently achieves the tightest interval across choices of α.

From Figure 3(**b**), it is clear that coverage is solely determined by the significance level, with all methods yielding nearly identical coverage across the entire range. The coverage probability of a confidence interval is defined as $(1-\alpha)\times 100\%$, and this relationship is validated by the downward sloping linear graph.

3.4 Null Hypothesis

Finally, we conclude the results section by evaluating each method's likelihood of committing a Type II error as the null hypothesis varies. In Figures 4 and 5, we fix the true variance at 15 and calculate the percentage of confidence intervals that result in a Type II error for each method. Figure 4 sets the null hypothesis to 10 (below the true variance), while Figure 5 sets the null hypothesis to 20 (above the true variance).

In Figure 4, where the null hypothesis states the population variance is smaller than it truly is, 95 out of 100 optimal confidence intervals incorrectly failed to reject the null hypothesis. For the equal area tail confidence interval, 83 out of the 100 equal area tail confidence intervals incorrectly

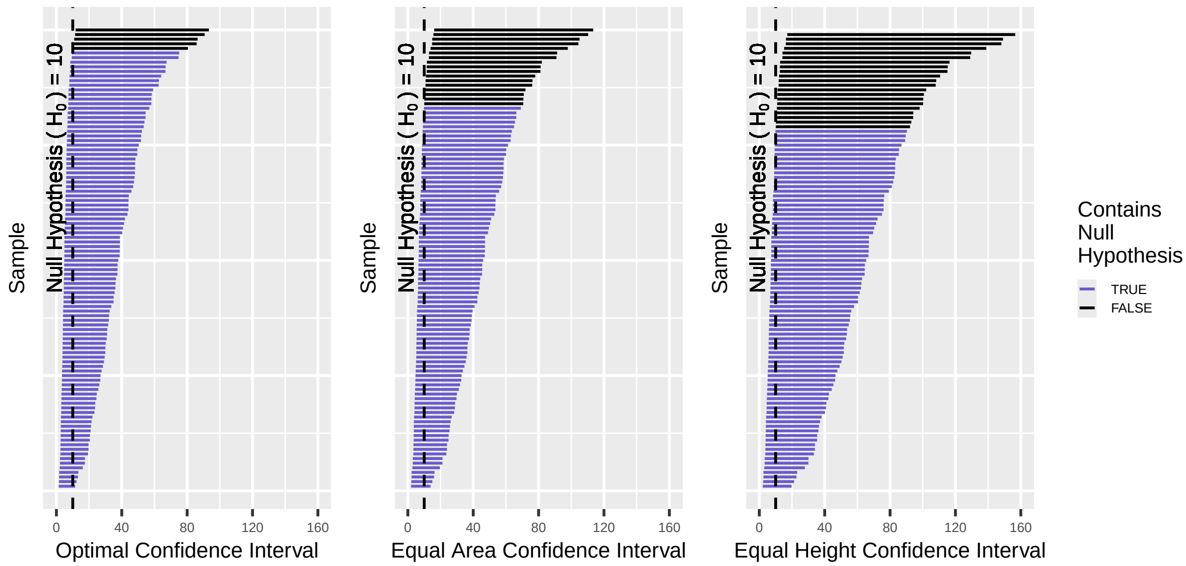


Figure 4: Performance of 95% Confidence Intervals in Accepting the Null Hypothesis ($H_0 = 10$) When the True Variance is $\sigma^2_{\text{true}} = 15$ ($m = 100, n = 10$). The inclusion of the null hypothesis, which is smaller than the true variance, indicates a Type II error.

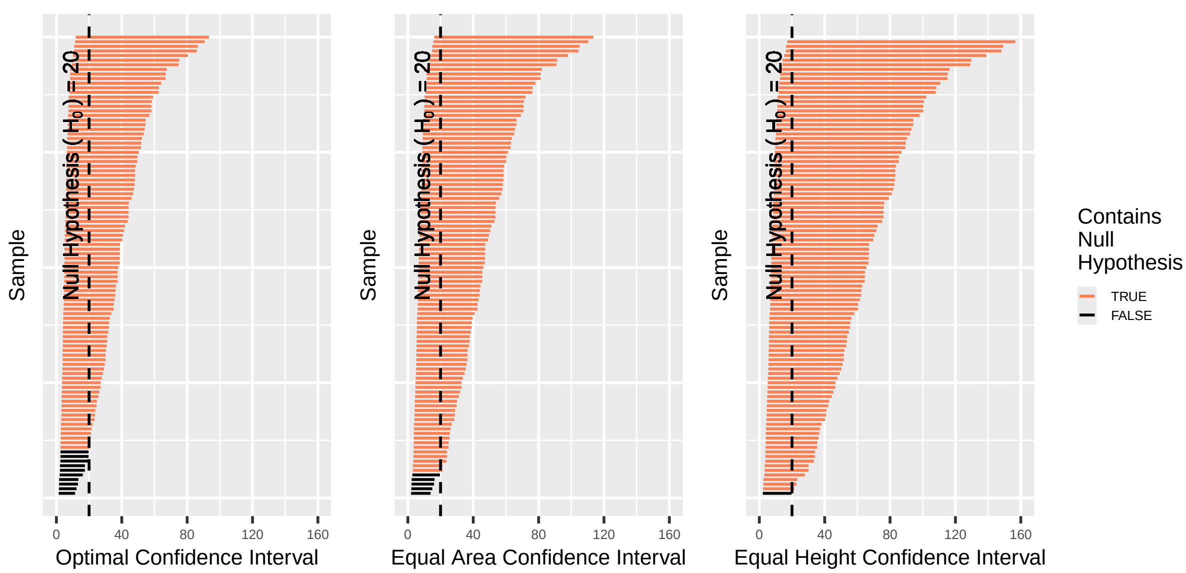


Figure 5: Performance of 95% Confidence Intervals in Accepting the Null Hypothesis ($H_0 = 20$) When the True Variance is $\sigma^2_{\text{true}} = 15$ ($m = 100, n = 10$). The inclusion of the null hypothesis, which is larger than the true variance, indicates a Type II error.

failed to reject the null hypothesis. And finally, for the equal height tail confidence interval, 79 out of the 100 equal height tail confidence intervals incorrectly failed to reject the null hypothesis. The optimal confidence interval committed the most Type II errors while the equal height confidence interval committed the least.

In Figure 5, where the null hypothesis states the population variance is larger than it truly is, 90 out of 100 optimal confidence intervals incorrectly failed to reject the null hypothesis. For the equal area tail confidence interval, 95 out of the 100 equal area tail confidence intervals incorrectly failed to reject the null hypothesis. And finally, for the equal height tail confidence interval, 99 out of the 100 equal height tail confidence intervals incorrectly failed to reject the null hypothesis. The optimal confidence interval committed the least Type II errors while the equal height confidence interval committed the most.

It seems a little strange that there isn't any consistency between the two figures, so we check the Type II error rates across multiple null hypotheses, not just the 2 shown from Figures 4 and 5.

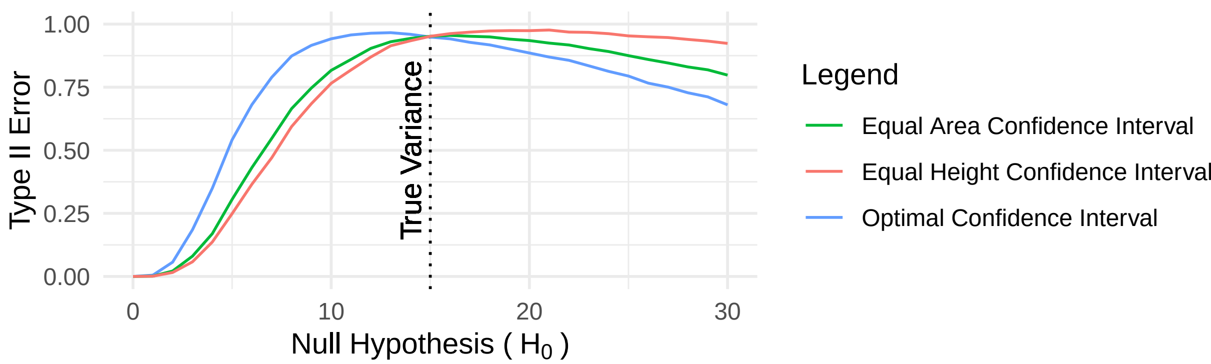


Figure 6: Type II Error Trends Across Null Hypotheses (H_0) for Confidence Intervals $\left(m = 10000, n = 10, \sigma^2_{\text{true}} = 15, \alpha = 0.05\right)$

In Figure 6, the equal height tail confidence interval shows the lowest Type II error rate when the null hypothesis is less than the true variance, but the highest Type II error rate when the null hypothesis is greater than the true variance. In contrast, the optimal confidence interval exhibits the opposite pattern, with the highest Type II error rate when the null hypothesis is less than the true variance and the lowest Type II error rate when the null hypothesis exceeds the true variance.

The reason for this behavior lies in the fact that the critical values for the optimal confidence interval are located at quantiles further to the right compared to those of the equal area tail and equal height tail confidence intervals. Specifically, for the optimal confidence interval, the lower and upper critical values satisfy the conditions $q_1' > \alpha/2$ and $q_2' > 1 - \alpha/2$. The confidence interval width is given by $\left[\frac{(n-1)s^2}{\chi^2_{\text{Upper}}}, \frac{(n-1)s^2}{\chi^2_{\text{Lower}}}\right]$, which for the optimal confidence interval becomes

$$\left[\frac{(n-1)s^2}{\texttt{qchisq(}q_2'\texttt{, n - 1)}}, \frac{(n-1)s^2}{\texttt{qchisq(}q_1'\texttt{, n - 1)}}\right],$$

in `R`. For the equal area tail confidence interval, it is

$$\left[\frac{(n-1)s^2}{\texttt{qchisq(}1-\alpha/2\texttt{, n - 1)}}, \frac{(n-1)s^2}{\texttt{qchisq(}\alpha/2\texttt{, n - 1)}}\right].$$

Since $q'_1 > \alpha/2$ and $q'_2 > 1-\alpha/2$, it follows that

$$\frac{(n-1)s^2}{\texttt{qchisq}(q'_2\texttt{, n - 1)}} < \frac{(n-1)s^2}{\texttt{qchisq}(1-\alpha/2\texttt{, n - 1)}} \text{ and } \frac{(n-1)s^2}{\texttt{qchisq}(q'_1\texttt{, n - 1)}} < \frac{(n-1)s^2}{\texttt{qchisq}(\alpha/2\texttt{, n - 1)}}.$$

Thus, the optimal confidence interval yields both smaller lower and upper bounds compared to the equal area tail interval. This explains why the optimal interval results in a higher Type II error when the null hypothesis assumes the true variance is smaller than it actually is.

Since the optimal confidence interval is less centered and shifted to the left, it tends to be more tolerant of small sample variances and null hypotheses, making it more likely to fail to reject them. Even though the optimal interval has the smallest width overall, this shift causes it to fail to reject null hypotheses that the equal area confidence interval would reject. In particular, for a null hypothesis between the quantiles $\alpha/2$ and q'_1, the optimal confidence interval is more likely to accept incorrect null hypotheses.

Conversely, due to the asymmetry of the chi-square distribution, the quantiles for the equal height tail method are positioned even further to the left than the equal area tail method, leading it to commit more Type II errors than the other methods when the null hypothesis suggests a variance greater than the true value.

4 Conclusion

In this study, we evaluated the performance of three confidence interval methods — equal area tail, equal height tail, and optimal — across various sample sizes, significance levels, and true variances. Each method was assessed on key metrics including interval width, coverage probability, and Type II error rates. Based on the results, we can make several recommendations on which methods are ideal for specific scenarios.

For small sample sizes, the optimal confidence interval method consistently yields the smallest interval widths and thus provides the most precise estimates. However, as sample size increases, the differences in width between the methods become less pronounced, and all methods converge, with the equal area tail method offering a reasonable balance between accuracy and computational simplicity.

When considering the true variance, the optimal method remains the most efficient across the range of variances tested, producing consistently smaller intervals compared to both the equal area tail and equal height tail methods. The equal height tail method, while producing wider intervals, shows a tendency to have the lowest Type II error rate when the null hypothesis is less than the true variance, making it a preferable choice when minimizing Type II errors is a priority in such scenarios.

Finally, when assessing Type II error rates as the null hypothesis varies, we observed that the equal height tail method performs better when the null hypothesis is below the true variance, whereas the optimal method excels when the null hypothesis is above the true variance, offering the lowest Type II error rate in that case.

In summary:

1. For scenarios where precise estimates are required and sample sizes are small, the optimal method is ideal.

2. For larger sample sizes, the equal area tail method offers a good trade-off between simplicity and performance.
3. For minimizing Type II errors when the null hypothesis is greater than the true variance, the optimal method is preferred.
4. For minimizing Type II errors when the null hypothesis is less than the true variance, the equal height tail method is recommended.

Thus, the choice of method depends largely on the specific context of the analysis, such as sample size, the true variance, and the trade-off between interval width and error rates.

The findings of this study provide valuable insights into the performance of confidence interval methods under varying conditions, but several limitations should be acknowledged. First, the study assumes the underlying data follow a normal distribution. While this assumption is common in theoretical evaluations, it may not always hold in real-world applications where data distributions may be skewed or heavy-tailed. Future research should extend these methods to non-normal distributions to assess their robustness in broader contexts.

Second, the study evaluates performance based on specific metrics such as interval width, coverage probability, and Type II error rates. While these metrics are critical, other considerations, such as computational efficiency and ease of implementation, were not explored in detail. These factors can significantly influence the practical applicability of the methods in large-scale or real-time analyses.

Additionally, the research was conducted under a controlled set of conditions, including specific sample sizes, significance levels, and true variances. Expanding the study to include a wider range of conditions, such as extreme variances or highly unbalanced sample sizes, would provide a more comprehensive evaluation.

Future research should investigate the underlying reasons why the optimal confidence interval tends to produce bounds that consistently favor reducing both the lower and upper limits, rather than the reverse. This exploration would involve examining the behavior of the chi-square distribution across different conditions and understanding how this pattern emerges systematically.

In conclusion, while the study offers practical recommendations for selecting confidence interval methods in various scenarios, there remains significant potential for further exploration. Addressing the limitations outlined above and pursuing new directions in confidence interval research will enhance the applicability and robustness of these methods in statistical practice.

Appendix

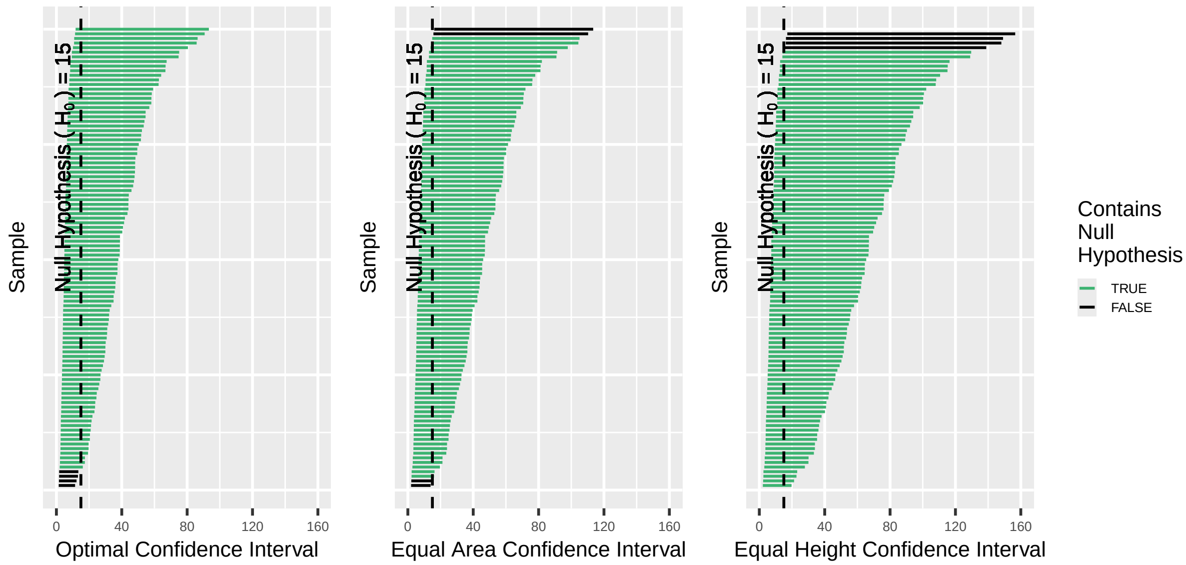


Figure 7: Performance of 95% Confidence Intervals in Accepting the Null Hypothesis ($H_0 = 15$) When the True Variance is $\sigma^2_{\text{true}} = 15$ ($m = 100, n = 10$). The exclusion of the null hypothesis, which is equal to the true variance, signifies a Type I error.

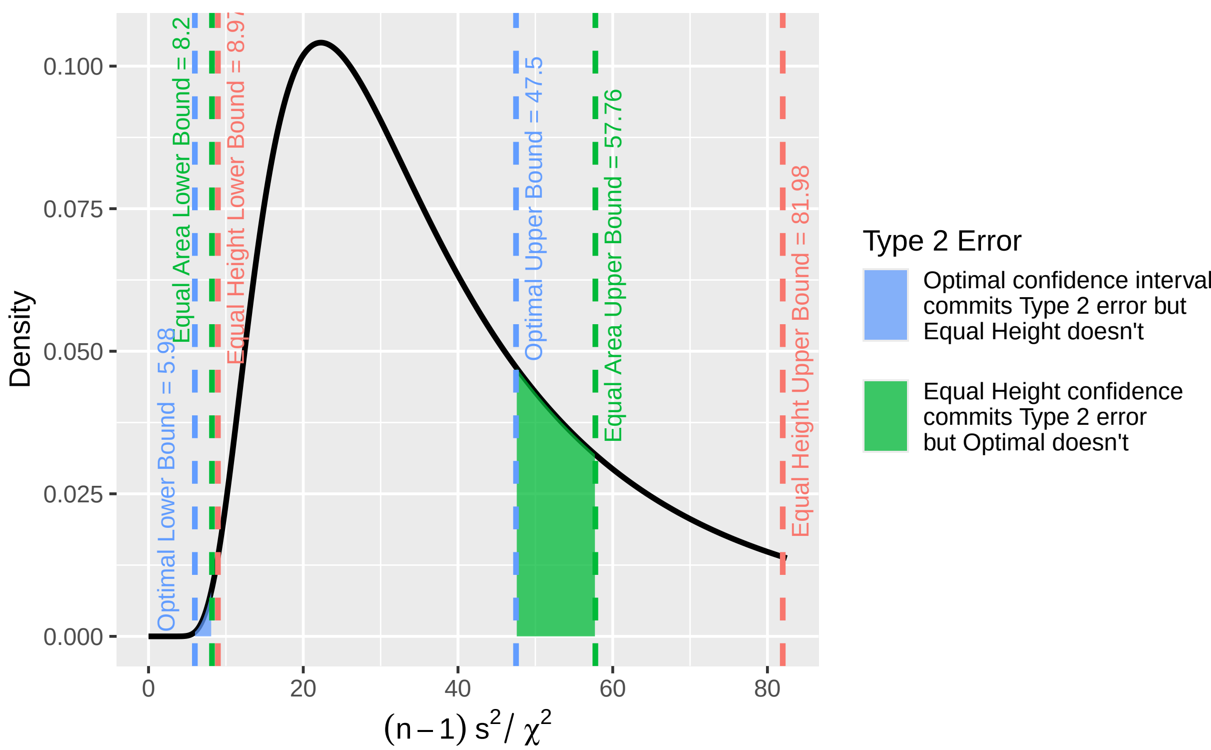


Figure 8: Comparison of the 95% Optimal, Equal Area, and Equal Height Confidence Intervals for $n = 10$, $\sigma^2_{\text{true}} = 15$, and $\alpha = 0.05$

References

Cojbasic, Vesna, and Dragan Loncar. "One-sided confidence intervals for population variances of skewed distributions." Journal of Statistical Planning and Inference, vol. 141, no. 5, May 2011, pp. 1667–1672, https://doi.org/10.1016/j.jspi.2010.11.007.

Shorrock, Glenn. "Improved confidence intervals for a normal variance." The Annals of Statistics, vol. 18, no. 2, 1 June 1990, https://doi.org/10.1214/aos/1176347636.

Tate, R. F., and G. W. Klett. "Optimal confidence intervals for the variance of a normal distribution." Journal of the American Statistical Association, vol. 54, no. 287, Sept. 1959, pp. 674–682, https://doi.org/10.1080/01621459.1959.10501528.