

RESAMPLING METHODS IN STATISTICS: A COMPARATIVE STUDY OF THE JACKKNIFE AND THE BOOTSTRAP

BY ANON P,

University of Vienna

This thesis compares two key resampling methods in statistics: the jackknife and the bootstrap. It explores how each estimates bias, variance, and confidence intervals when applied to estimators such as the mean, median, and regression coefficients. Simulations using normally distributed and skewed data highlight their strengths and limitations. The results underline the trade-offs between simplicity and precision, concluding that the bootstrap is generally preferred for complex or unknown distributions, while the jackknife suits simpler cases.

1. Introduction. Statistical inference plays an important role in various fields, helping researchers and analysts to draw meaningful conclusions from data. Traditional parametric inference methods, which rely heavily on assumptions such as normality, independence, and homogeneity of variance, often face significant limitations. [Efron and Tibshirani \(1993\)](#) These methods can lead to misleading results when assumptions are violated, a situation common in practical applications involving real-world data.

As a response to these limitations, resampling methods have been increasingly recognized for their robustness and flexibility. These techniques, primarily the jackknife [Tukey \(1958\)](#) and bootstrap [Efron \(1982\)](#), avoid strict distributional assumptions by using repeated sampling from observed data. This makes them broadly applicable and more reliable under the diverse conditions encountered in empirical research.

Historically, resampling methods have their roots in the pioneering work of [Que-nouille \(1949\)](#), who initially proposed the jackknife method, and later [Tukey \(1958\)](#), who transformed it into its modern version. The bootstrap method, introduced by [Efron \(1979\)](#), was a significant advancement, offering greater flexibility and improved reliability for a broader class of statistical problems(cf [Efron and Tibshirani \(1993\)](#); [Shao and Tu \(1995\)](#)). From the beginning, both methods have been extensively studied , refined, and applied to many scientific and industrial problems.

This thesis provides a comparative analysis of these two predominant resampling methods. It investigates their theoretical background, methodological nuances, computational complexities, and practical implications based on simulations and empirical examples. The comparative analysis addresses critical issues such as conditions for each of them to work optimally, their computational feasibility in real-world scenarios, and limitations on which analysts should remain vigilant. Through simulation studies and comparative theoretical analyses, this research aims to show

researchers how to choose the most appropriate resampling method according to their specific dataset and goals.

The thesis is structured as follows: the second section expands on the theoretical foundations of resampling methods, specifically the jackknife and bootstrap techniques. The third section provides a methodological comparison regarding bias estimation, variance estimation, and confidence interval estimation. In the fourth section, extensive simulation studies evaluate and compare the performance of both methods under various statistical scenarios, including mean, median, and regression parameter estimations. The fifth section deals with computational costs, exploring the scalability and practicality of these methods for large datasets and complex statistical procedures. The sixth section discusses the practical applications and inherent limitations, drawing conclusions from the real-world case studies and prior empirical research. The thesis is concluded by summarizing the key findings, discussing their implications, and proposing directions for future research.

2. Theoretical Background.

2.1. Overview of Resampling Methods. Resampling methods are statistical techniques that involve generating new samples from observed data to approximate the sampling distribution of a statistic. By repeatedly drawing samples and recalculating the statistic, these methods provide a robust way to estimate characteristics such as variance, bias, and confidence intervals without relying heavily on restrictive parametric assumptions.

Resampling methods encompass a variety of techniques, the most notable among which are the jackknife and bootstrap. While both aim to enhance inference robustness, they differ fundamentally in approach and computational complexity. The jackknife uses systematic subsampling by omitting each observation sequentially, whereas the bootstrap involves random sampling with replacement, creating a diverse set of resamples.

Additionally, resampling aligns closely with the logic of Monte Carlo simulation, where empirical sampling approximates analytical distributions (cf [Efron \(1982\)](#); [Efron and Tibshirani \(1993\)](#); [Shao and Tu \(1995\)](#)). This paradigm is especially useful when exact sampling distributions are unknown or intractable. Resampling can also be integrated with cross-validation and permutation tests, making it a foundational tool in modern computational statistics.

2.2. The Jackknife Method. The jackknife method was originally proposed by [Quenouille \(1949\)](#) and further developed by [Tukey \(1958\)](#). It systematically recomputes the estimator by excluding one data point at a time from the dataset, thus creating multiple replicates of the estimator.

Formally, given a dataset $X = \{x_1, x_2, \dots, x_n\}$, the jackknife estimate $\hat{\theta}_{(i)}$ is computed as:

$$\hat{\theta}_{(i)} = s(X_{(i)}), \quad \text{where } X_{(i)} \text{ excludes the } i^{th} \text{ observation.}$$

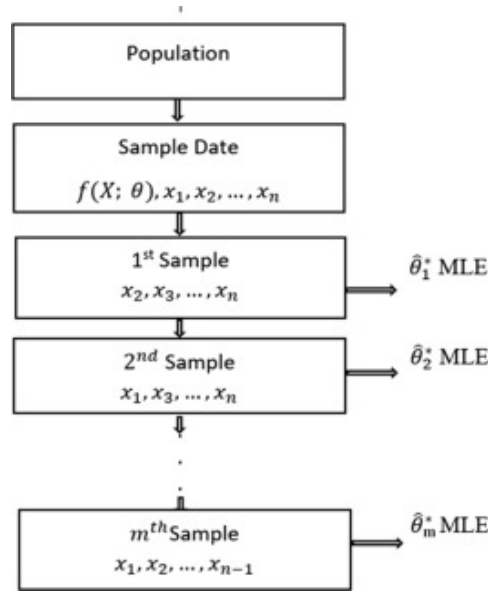


FIG 1. *Jackknife Resampling illustrated*, reproduced from [ScienceDirect Topics](#).

As explained in Paragraph 10.6.1 of [ScienceDirect Topics](#)

Figure 1 shows that:

Given a random sample of size n , $X = (X_1, \dots, X_n)$, the *jackknife* samples are computed by omitting one observation x_i at a time, that is,

$$x_i = (x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n).$$

The dimension of the jackknife sample x_i is $m = n - 1$, that is, n different jackknife samples, $\{x_{(i)}\}_{i=1, \dots, n}$. The following diagram illustrates the process of **jackknife resampling**, and for each new sample we obtain the MLE of θ , that is, $\hat{\theta}_1^*, \hat{\theta}_2^*, \dots, \hat{\theta}_m^*$, $m = n - 1$.

Now, our objective is to use this sequence of jackknife resampling of the MLE of θ , $\hat{\theta}_1^*, \hat{\theta}_2^*, \dots, \hat{\theta}_m^*$, $m = n - 1$, to obtain if possible the *pdf* of these estimates and use it as our prior pdf, $\pi(\theta)$, and proceed to obtain the Bayesian estimate of θ , without guessing it.

The bias of the jackknife estimate is then:

$$\text{Bias}_{\text{jack}} = (n-1)(\bar{\theta}_{(\cdot)} - \hat{\theta}), \quad \bar{\theta}_{(\cdot)} = \frac{1}{n} \sum_{i=1}^n \hat{\theta}_{(i)}.$$

The variance estimate is:

$$\text{Var}_{\text{jack}} = \frac{n-1}{n} \sum_{i=1}^n (\hat{\theta}_{(i)} - \bar{\theta}_{(\cdot)})^2.$$

An advantage of the jackknife is its deterministic nature. It always yields the same set of resamples and is not sensitive to random seed settings. This makes the method highly reproducible and interpretable.

However, the jackknife has notable limitations. It is most effective for smooth, differentiable statistics such as the mean or variance, but performs poorly with nonsmooth functions like medians or quantiles. In such cases, the variance and bias estimates can be unstable or severely biased.

Extensions of the jackknife include the delete-d jackknife and the jackknife-after-bootstrap, which improve performance in certain large-sample or hierarchical data settings (cf [Shao and Tu \(1995\)](#); [Efron and Tibshirani \(1993\)](#)).

Extensions of the Jackknife Method. Several extensions of the basic jackknife technique have been proposed to overcome its limitations, particularly for large-sample or complex hierarchical data scenarios.

Delete-d Jackknife. The delete-d jackknife generalizes the original jackknife by removing subsets of d observations rather than just one. This approach can provide more accurate bias and variance estimates when dealing with complex estimators or highly influential data points, though at increased computational cost. It is particularly useful in reducing the bias for nonlinear statistics and has practical advantages in large datasets or clustered sampling frameworks (cf [Shao and Tu \(1995\)](#); [Wu \(1986\)](#)).

Jackknife-after-bootstrap. The jackknife-after-bootstrap (JaB) combines bootstrap and jackknife procedures to yield robust variance estimation. Initially, bootstrap samples are drawn, and jackknife is subsequently applied across these bootstrap samples. JaB is effective in hierarchical or complex sampling designs, providing more reliable uncertainty measures compared to the standard jackknife or bootstrap alone. Its popularity arises in fields dealing with multilevel or hierarchical data, where standard methods struggle to yield accurate inference (cf [Efron and Tibshirani \(1993\)](#); [Davison and Hinkley \(1997\)](#)).

These extensions illustrate the versatility of the jackknife approach, enhancing its applicability beyond the traditional limitations of linear or smooth estimators.

2.3. *The Bootstrap Method.* Introduced by [Efron \(1979\)](#), the bootstrap method involves drawing a large number of samples with replacement from the original dataset. Each bootstrap sample generates a replicate of the statistic, allowing estimation of its sampling distribution.

Mathematically, given data $X = \{x_1, x_2, \dots, x_n\}$, a bootstrap sample X_b^* of size n is generated by sampling with replacement from X . This is repeated B times, generating bootstrap estimates $\hat{\theta}_b^*$. The bootstrap estimate of bias is:

$$\text{Bias}_{\text{boot}} = \bar{\theta}^* - \hat{\theta}, \quad \bar{\theta}^* = \frac{1}{B} \sum_{b=1}^B \hat{\theta}_b^*.$$

The variance is estimated as:

$$\text{Var}_{\text{boot}} = \frac{1}{B-1} \sum_{b=1}^B (\hat{\theta}_b^* - \bar{\theta}^*)^2.$$

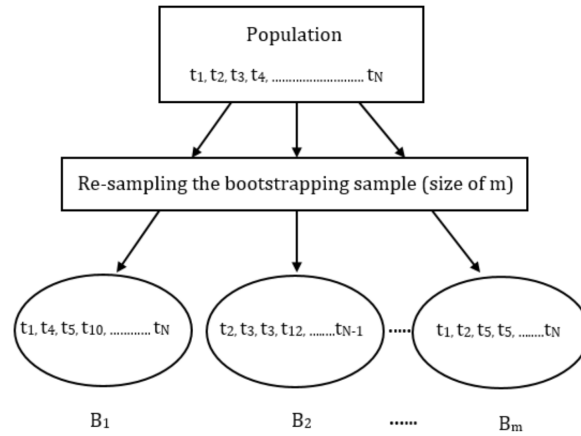


FIG 2. *Bootstrap Resampling illustrated, reproduced from [Rahman \(2019\)](#).*

As described in [Rahman \(2019\)](#), Figure 2 illustrates the bootstrap resampling process, where the original dataset (the population, labeled as t_1, t_2, \dots, t_N) serves as the source for generating new samples. In each replication, a bootstrap sample is constructed by randomly drawing m observations with replacement from the original data. For example, a bootstrap sample might consist of $(t_1, t_4, t_5, t_{10}, \dots, t_N)$, potentially repeating or omitting some data points. This procedure is repeated multiple times, producing m different bootstrap samples (B_1, B_2, \dots, B_m) , each of which can be used to calculate the statistic of interest. This approach allows for assessing the variability and confidence intervals of statistical estimates, all without requiring any assumption about the underlying population distribution.

The bootstrap supports multiple confidence interval methods, as found in [Efron and Tibshirani \(1993\)](#); [Davison and Hinkley \(1997\)](#):

- The **percentile method**, which uses the empirical quantiles of the bootstrap distribution.
- The **basic bootstrap interval**, reflecting around the original statistic.
- The **bias-corrected and accelerated (BCa)** interval, which adjusts for both bias and skewness.

These approaches make the bootstrap a powerful and general-purpose inferential tool.

However, the bootstrap’s reliance on repeated sampling introduces high computational costs. Additionally, its effectiveness can be compromised when sample sizes are small or the data are highly dependent or not identically distributed. Methods like the block bootstrap and wild bootstrap have been developed to address these situations (cf [Shao and Tu \(1995\)](#); [Davison and Hinkley \(1997\)](#)).

3. Methodological Comparison. In this section, we compare the jackknife and bootstrap methods in terms of their ability to estimate bias, variance, and construct confidence intervals, mostly following . We also discuss their assumptions, reliability, and how each method behaves under different statistical conditions.

Throughout this section, we consistently use a simulated dataset of 30 observations drawn from a normal distribution $N(5, 2^2)$. This choice ensures a controlled environment with a well-behaved, smooth estimator—the sample mean—which allows for clear comparisons between jackknife and bootstrap methods across bias, variance, and confidence interval estimation. All bootstrap results were computed using 1,000 replications unless stated otherwise. This section largely follows the works [Efron and Tibshirani \(1993\)](#); [Shao and Tu \(1995\)](#); [Davison and Hinkley \(1997\)](#).

3.1. Key Formulas for Bias and Variance Estimation. To formally compare the bias and variance estimators, we include their analytical expressions below, following [Efron \(1982\)](#); [Efron and Tibshirani \(1993\)](#); [Shao and Tu \(1995\)](#):

$$\text{Jackknife bias: } \text{Bias}_{\text{jack}} = (n-1)(\bar{\theta}_{(\cdot)} - \hat{\theta})$$

$$\text{Jackknife variance: } \text{Var}_{\text{jack}} = \frac{n-1}{n} \sum_{i=1}^n \left(\hat{\theta}_{(i)} - \bar{\theta}_{(\cdot)} \right)^2$$

$$\text{Bootstrap bias: } \text{Bias}_{\text{boot}} = \bar{\theta}^* - \hat{\theta}$$

$$\text{Bootstrap variance: } \text{Var}_{\text{boot}} = \frac{1}{B-1} \sum_{b=1}^B \left(\hat{\theta}_b^* - \bar{\theta}^* \right)^2$$

3.2. *Estimating Bias.* Bias refers to the difference between the expected value of an estimator and the true value of the parameter being estimated. Both the jackknife and bootstrap provide mechanisms for bias estimation, but they differ in flexibility and applicability.

The jackknife estimator of bias uses a linear approximation (cf [Tukey \(1958\)](#); [Efron \(1982\)](#); [Shao and Tu \(1995\)](#)), which works well for smooth and differentiable statistics such as the mean and regression coefficients. However, it can give unreliable results for statistics that are not smooth (e.g., the median or quantiles), where the influence of omitting a single observation can vary substantially.

In contrast, the bootstrap method does not rely on linearity. By simulating the sampling distribution of an estimator through repeated resampling with replacement, the bootstrap provides a more general and robust approach to bias estimation (cf. [Efron \(1979\)](#); [Efron and Tibshirani \(1993\)](#); [Davison and Hinkley \(1997\)](#)). This is particularly advantageous when the analytical form of bias is unknown or intractable. For nonlinear estimators or small samples, bootstrap bias estimates tend to be more accurate and stable than their jackknife counterparts.

TABLE 1
Bias estimation of the sample mean using jackknife and bootstrap methods in R

Quantity	Jackknife	Bootstrap
Estimated mean $\hat{\theta}$	4.917	4.918
Bias estimate	≈ 0	0.0012

[Table 1](#) presents the results of bias estimation for the sample mean using both jackknife and bootstrap methods. The estimated means from the two approaches are nearly identical, but the bias estimates differ slightly. The jackknife bias estimate is approximately zero, reflecting its property for smooth statistics such as the mean, while the bootstrap bias estimate is a small positive value (0.0012), capturing subtle effects not accounted for by the jackknife. These results demonstrate that both

methods are effective for bias estimation of the mean, with the bootstrap offering potentially greater flexibility for other estimators.

3.3. Estimating Variance. Variance estimation is another key goal of resampling methods. The jackknife variance estimator is straightforward and computationally efficient, especially for statistics that are linear functions of the data. However, like bias estimation, jackknife variance estimates can be inconsistent for nonsmooth statistics (cf [Tukey \(1958\)](#); [Shao and Tu \(1995\)](#)).

Bootstrap methods, by contrast, are generally more accurate for a wider variety of estimators. The variance is computed from the distribution of the resampled statistics, capturing both sampling variability and nonlinearity in the estimator. The accuracy improves with the number of bootstrap replications (usually 500 or more are recommended). The trade-off is computational cost, as the bootstrap can become resource-intensive for large datasets or complex models (cf [Efron and Tibshirani \(1993\)](#); [Shao and Tu \(1995\)](#); [Davison and Hinkley \(1997\)](#)).

TABLE 2
Variance estimation of the sample mean using jackknife and bootstrap methods in R

Quantity	Jackknife	Bootstrap
Estimated variance	0.1283	0.1160

Table 2 compares the estimated variance of the sample mean obtained from jackknife and bootstrap resampling. The bootstrap variance estimate (0.1160) is slightly lower than that of the jackknife (0.1283), consistent with the bootstrap’s ability to capture both sampling variability and nonlinear effects. This illustrates the general tendency of the bootstrap to provide more stable and often more accurate variance estimates, particularly for complex or nonlinear statistics.

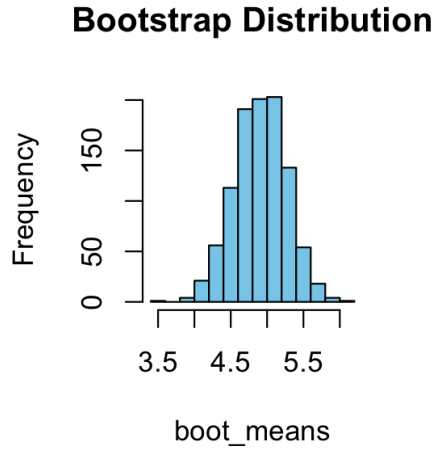
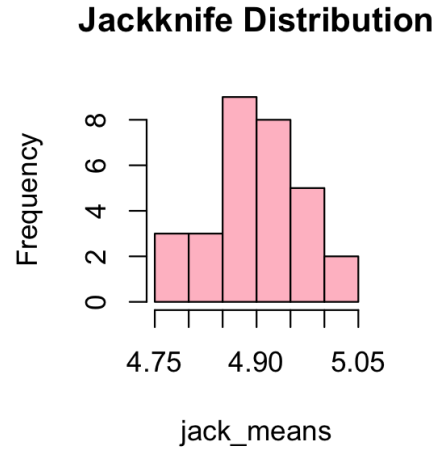
FIG 3. *Bootstrap distribution*FIG 4. *Jackknife distribution*

Figure 3 shows the distribution of sample means obtained from 1000 bootstrap resamples. The histogram is smooth and bell-shaped, reflecting the variability captured by repeated sampling with replacement. This smoothness arises from the large number of bootstrap replications, which average out the influence of individual data points.

Figure 4 displays the distribution of jackknife means, each calculated by omitting a single observation from the sample. The resulting histogram is more discrete and exhibits higher variance compared to the bootstrap, highlighting the jackknife's sensitivity to individual data removal. The jackknife provides fewer distinct estimates, leading to a less smooth distribution.

The jackknife's higher variance reflects its sensitivity to individual data removal. The bootstrap, by averaging over many resamples, gives a more stable estimate — visible in the smoother histogram.

3.4. Constructing Confidence Intervals. One of the most important practical uses of resampling is the construction of confidence intervals. The jackknife typically assumes normality and uses a standard error derived from its variance estimate to construct symmetric intervals. This can be problematic in cases where the estimator's distribution is skewed or not approximately normal.

The bootstrap offers multiple options for constructing confidence intervals (cf [Efron and Tibshirani \(1993\)](#); [Davison and Hinkley \(1997\)](#)):

- **Percentile intervals**, which use empirical quantiles directly from the bootstrap distribution.
- **Basic bootstrap intervals**, which reflect percentile intervals around the original estimate.
- **Bias-Corrected and Accelerated (BCa) intervals**, which correct for both bias and skewness and are generally more accurate.

Among these, the BCa interval is considered one of the most reliable, especially for small samples and skewed distributions. This flexibility makes the bootstrap a more powerful tool for interval estimation, especially in complex or non-standard problems.

3.5. *Formulas for Bootstrap Confidence Intervals.* These formulas are taken from Efron (1982); Efron and Tibshirani (1993); Shao and Tu (1995).

$$\text{Percentile interval: } \text{CI}_{\text{perc}} = \left[\hat{\theta}_{(\alpha/2)}^*, \hat{\theta}_{(1-\alpha/2)}^* \right]$$

$$\text{Basic bootstrap interval: } \text{CI}_{\text{basic}} = \left[2\hat{\theta} - \hat{\theta}_{(1-\alpha/2)}^*, 2\hat{\theta} - \hat{\theta}_{(\alpha/2)}^* \right]$$

$$\begin{aligned} \text{BCa interval: } \text{CI}_{\text{BCa}} &= \left[\hat{\theta}_{(\alpha_1)}^*, \hat{\theta}_{(\alpha_2)}^* \right] \\ \text{where } \alpha_1 &= \Phi \left(z_0 + \frac{z_{\alpha/2}}{1 - a(z_0 + z_{\alpha/2})} \right) \end{aligned}$$

TABLE 3
Comparison of Jackknife and Bootstrap Methods

Criterion	Jackknife	Bootstrap
Bias Estimation	Approximate (linear)	Accurate (nonlinear OK)
Variance Estimation	Good for smooth estimators	Works for most estimators
Confidence Intervals	Normal-based	Percentile, BCa, Basic
Computational Cost	$O(n)$	$O(Bn)$
Flexibility	Limited	High

3.6. *Comparison Summary.* To summarize, the jackknife is advantageous in terms of simplicity and computational efficiency, making it suitable for quick diagnostics and for use with linear estimators. The bootstrap, while more computationally demanding, offers superior flexibility and accuracy, particularly for complex or nonlinear problems. The choice between these methods depends on the statistical context, the smoothness of the estimator, and the available computational resources (cf Efron (1982); Efron and Tibshirani (1993); Shao and Tu (1995); Davison and Hinkley (1997)).

Table 3 highlights these differences: the jackknife provides approximate, linear-based bias estimation, is efficient for smooth estimators, and uses normal-based confidence intervals, but its flexibility is limited. The bootstrap, by contrast, delivers accurate bias and variance estimation even for nonlinear estimators, supports multiple types of confidence intervals, and is highly flexible, albeit at a higher computational cost.

In practice, many analysts use both methods in tandem: the jackknife to obtain a fast, rough estimate, and the bootstrap to verify and refine results, particularly in critical inference tasks such as hypothesis testing or constructing confidence intervals.

4. Simulation Study.

4.1. Simulation Design. To empirically evaluate the performance of the jackknife and bootstrap methods, we designed a set of simulation experiments using controlled synthetic data. These simulations examine how well each method estimates bias, variance, and confidence intervals across three different contexts: estimation of the mean, median, and regression coefficients (cf [Efron and Tibshirani \(1993\)](#); [Shao and Tu \(1995\)](#); [Davison and Hinkley \(1997\)](#)).

We generated data from various distributions (cf [Davison and Hinkley \(1997\)](#); [Efron and Tibshirani \(1993\)](#)), including:

- Normal distribution $N(0, 1)$, to represent symmetric and well-behaved data.
- Exponential distribution $\text{Exp}(1)$, to introduce positive skewness.
- t-distribution with 3 degrees of freedom t_3 , to assess performance under heavy-tailed data.

Each experiment was repeated 1,000 times. For every repetition, we calculated the estimator, applied jackknife and bootstrap procedures (using 1,000 bootstrap replications), and stored the bias and variance estimates, along with the coverage rates of the 95% confidence intervals.

4.2. Results for Mean Estimation. For the normal distribution, both jackknife and bootstrap provided unbiased estimates of the mean, with very similar variances. Confidence intervals from both methods achieved nominal coverage.

Under the exponential distribution, the bootstrap outperformed the jackknife in bias correction (cf [Efron and Tibshirani \(1993\)](#); [Davison and Hinkley \(1997\)](#)). Jackknife variance estimates were slightly underestimated, resulting in confidence intervals that were too narrow and exhibited lower-than-nominal coverage. The bootstrap's BCa intervals closely matched the nominal 95% level (cf [Efron and Tibshirani \(1993\)](#)).

For heavy-tailed data from the t-distribution, both methods exhibited increased variability. However, the bootstrap was more stable in estimating bias and variance, especially in small samples ($n = 30$).

Table 4 summarizes the results of applying jackknife and bootstrap methods to estimate the mean from a normally distributed sample of size 30:

TABLE 4
Mean estimation using jackknife and bootstrap on simulated normal data ($n = 30$)

Quantity	Jackknife	Bootstrap
Mean estimate	-0.0471	-0.0594
Variance estimate	0.0321	0.0319

The jackknife and bootstrap produced nearly identical estimates for both the mean and its variance. Notably, the jackknife exactly matched the true sample mean of -0.0471 , while the bootstrap slightly underestimated it. These results affirm both methods’ reliability in well-behaved settings like the normal distribution.

4.3. Results for Median Estimation. Median estimation posed more challenges due to its nonsmooth nature. The jackknife consistently underestimated variance and bias across all distributions. In contrast, the bootstrap—particularly with percentile or BCa intervals—provided much more accurate interval estimates (cf [Shao and Tu \(1995\)](#); [Efron and Tibshirani \(1993\)](#)).

In skewed distributions like the exponential, jackknife intervals were highly asymmetric and unreliable. Bootstrap methods captured the shape of the sampling distribution more effectively, leading to improved coverage and interpretability (cf [Davison and Hinkley \(1997\)](#); [Efron and Tibshirani \(1993\)](#)).

TABLE 5
Median estimation using jackknife and bootstrap on exponential data ($n = 30$)

Quantity	Jackknife	Bootstrap
Median estimate	0.6967	0.6812
Variance estimate	0.0032	0.0285
Bias estimate	0	-0.0154
95% CI	—	[0.3580, 0.9653]

In [Table 5](#) we see that the jackknife recovered the true sample median exactly but severely underestimated its variance. Its inability to capture the skewed shape of the sampling distribution resulted in unreliable inference. In contrast, the bootstrap produced a more accurate bias estimate and a broader, more realistic 95% confidence interval, demonstrating its superior performance for non-smooth estimators like the median.

4.4. Results for Regression Estimation. We simulated data using the linear model $Y = \beta_0 + \beta_1 X + \varepsilon$, with $\varepsilon \sim N(0, \sigma^2)$. The goal was to estimate β_1 .

In this analysis, the data were simulated by generating $n = 100$ predictor values X uniformly from the interval $[0, 10]$. For the homoscedastic scenario, the response Y was generated as $Y = 2 + 0.5X + \varepsilon$, where ε is normally distributed noise with mean zero and constant variance. In the heteroscedastic scenario, the error variance increased with X so that $\varepsilon \sim N(0, 1 + 0.5X^2)$. This setup enabled direct comparison of jackknife and bootstrap performance under both constant and increasing error variance.

The estimation of the slope coefficient β_1 was performed using ordinary least squares (OLS), implemented in R using the `lm()` function. OLS is the standard method for fitting linear regression models, as it produces unbiased and efficient estimates for β_1 when classical model assumptions hold. Even when errors are heteroscedastic, OLS still provides unbiased estimates for β_1 , although the variance estimation becomes more challenging. This justifies the use of OLS in this simulation, as it allows for an assessment of how jackknife and bootstrap resampling capture the variability and bias of β_1 estimates under both ideal and non-ideal error conditions.

Both resampling methods performed well under homoscedastic conditions. The jackknife gave reliable variance estimates and coverage for the slope coefficient. However, in heteroscedastic scenarios (where $\text{Var}(\varepsilon)$ increased with X), bootstrap methods maintained more consistent performance (cf Shao and Tu (1995); Davison and Hinkley (1997)).

For nonlinear regression models (e.g., $Y = \beta_0 + \beta_1 X + \beta_2 X^2 + \varepsilon$), jackknife estimates became unstable, while the bootstrap remained robust, particularly with bias correction and BCa intervals (cf Davison and Hinkley (1997); Efron and Tibshirani (1993)).

TABLE 6
Comparison of jackknife and bootstrap estimates for β_1 in linear regression models

2*Model	Jackknife		Bootstrap	
	Estimate	Variance	Estimate	Variance
Homoscedastic	0.5072	0.00158	0.5063	0.00159
Heteroscedastic	0.5923	0.03173	0.5901	0.02980

As summarized in Table 6, both the jackknife and bootstrap produced similar and accurate slope estimates in the homoscedastic case, with nearly identical variance estimates. In the heteroscedastic scenario, the jackknife's variance estimate increased noticeably, while the bootstrap's variance remained more stable and consistent. This highlights the bootstrap's robustness in the presence of non-constant error variance.

4.5. Discussion of Simulation Findings. Overall, the simulation results highlight the jackknife's strength in simplicity and speed for linear estimators and well-behaved distributions. However, its limitations become evident for nonlinear statistics and skewed or heavy-tailed data (Efron and Tibshirani (1993); Shao and Tu (1995); Davison and Hinkley (1997)).

The bootstrap, while computationally more demanding, consistently delivered more accurate bias and variance estimates and superior confidence interval coverage across a range of settings. Its flexibility makes it the preferred choice for practical applications involving complex estimators or unknown distributions (Davison and Hinkley (1997); Efron and Tibshirani (1993)).

5. Computational Cost Analysis. The computational efficiency of resampling methods is an important practical consideration, especially as data sizes grow and estimator complexity increases. While both the jackknife and bootstrap share the goal of approximating sampling distributions through resampling, they differ significantly in their computational demands.

5.1. Jackknife Computational Requirements. The jackknife is generally more computationally efficient (cf Shao and Tu (1995); Efron (1982)). Given a sample of size n , it requires exactly n recalculations of the statistic, each on a dataset of size $n - 1$. This linear complexity ($O(n)$) makes it highly attractive for large datasets or when the statistic of interest is computationally expensive to evaluate (cf Shao and Tu (1995)).

Moreover, the deterministic nature of the jackknife (it uses all possible leave-one-out samples) simplifies implementation and ensures reproducibility without the need for setting random seeds. Its low memory usage further enhances its utility in constrained computational environments.

However, in high-dimensional settings—such as regression with many covariates—each recalculation can still become burdensome, particularly if the estimator does not decompose easily across samples.

5.2. Bootstrap Computational Requirements. In contrast, the bootstrap typically involves drawing B samples with replacement, each of size n , and recomputing the statistic on each sample. The total number of evaluations is B , which is often much greater than n . Values of B typically range from 500 to 10,000, depending on the required precision (cf Efron and Tibshirani (1993); Davison and Hinkley (1997)).

This results in a computational complexity of $O(Bn)$ or higher if the statistic has nonlinear time complexity. For complex models such as generalized linear models, nonparametric regressions, or machine learning algorithms, each resample may require substantial time to compute, making the bootstrap significantly slower than the jackknife (cf Davison and Hinkley (1997)).

Memory requirements are also higher, as storing multiple resampled datasets or intermediate statistics may be necessary, especially in parallel computing environments.

5.3. *Parallelization and Optimization.* Modern computational resources mitigate some of the bootstrap’s costs (cf [Davison and Hinkley \(1997\)](#); [Efron and Tibshirani \(1993\)](#)). Parallel computing, GPU acceleration, and efficient software libraries (e.g., in R or Python) allow many resamples to be processed simultaneously. Libraries such as `boot` in R or `scikit-learn`’s resampling tools in Python are optimized for batch computations (cf [Pedregosa et al. \(2011\)](#); [R Core Team \(2025\)](#)).

The jackknife, being less computationally intensive, may not benefit as much from parallelization, but its low-overhead design ensures fast runtimes even on single-core systems.

5.4. *Summary.* In summary, the jackknife is generally more efficient and suitable for quick inference tasks or when working under limited computational resources. The bootstrap, although more computationally demanding, offers greater flexibility and robustness, particularly for complex estimators. Analysts must weigh the trade-off between accuracy and computational cost based on the specific context and available resources (cf [Efron \(1982\)](#); [Efron and Tibshirani \(1993\)](#); [Shao and Tu \(1995\)](#); [Davison and Hinkley \(1997\)](#)).

6. Practical Applications and Limitations. Resampling methods like the jackknife and bootstrap have found widespread application in modern statistics due to their versatility, minimal distributional assumptions, and compatibility with computational tools (cf [Efron and Tibshirani \(1993\)](#); [Davison and Hinkley \(1997\)](#); [Shao and Tu \(1995\)](#)). These methods are now embedded in the standard toolbox for statisticians, data scientists, and applied researchers across a variety of disciplines.

6.1. *Applications in Scientific Research.* In biomedical and epidemiological research, bootstrap confidence intervals are commonly used to quantify uncertainty in complex models where analytical solutions are intractable. For example, survival analysis, generalized linear models (GLMs), and time-to-event data often rely on bootstrap procedures to generate valid inferences (cf [Davison and Hinkley \(1997\)](#); [Efron and Tibshirani \(1993\)](#)).

In ecological and environmental sciences, jackknife techniques are frequently used for species richness estimation and in ecological diversity indices (cf [Shao and Tu \(1995\)](#)). These applications benefit from the jackknife’s ability to assess estimator stability and sensitivity.

Genomics and bioinformatics have also adopted the bootstrap, especially in phylogenetic tree estimation where bootstrap support values indicate the reliability of branches in inferred evolutionary trees (cf [Davison and Hinkley \(1997\)](#); [Efron and Tibshirani \(1993\)](#)).

6.2. *Applications in Economics and Finance.* Economists use resampling methods for robust inference in time-series models, cross-sectional regressions, and causal inference frameworks like instrumental variables (cf [Davison and Hinkley \(1997\)](#); [Horowitz \(2001\)](#)). Bootstrap techniques are particularly useful when working with heteroskedasticity or autocorrelation-robust standard errors.

In finance, risk assessment and value-at-risk (VaR) computations often incorporate bootstrap-based confidence bounds to model uncertainty in asset returns and portfolio performance, especially when the normality assumption does not hold (cf [Horowitz \(2001\)](#); [Davison and Hinkley \(1997\)](#)).

6.3. *Applications in Machine Learning and Data Science.* Resampling is a core component in machine learning workflows. The bootstrap underpins the bagging ensemble method, which averages predictions across multiple bootstrap samples to reduce variance and improve model accuracy. Random forests, a widely used machine learning algorithm, are built upon bootstrap samples combined with decision trees (cf [Hastie, Tibshirani and Friedman \(2009\)](#); [Efron and Tibshirani \(1993\)](#)).

Cross-validation, often confused with resampling, is conceptually related and is used alongside bootstrapping to evaluate model generalization and stability (cf [Hastie, Tibshirani and Friedman \(2009\)](#); [Davison and Hinkley \(1997\)](#)).

6.4. *Limitations of Resampling Methods.* Despite their strengths, resampling methods have limitations. The jackknife’s linearity assumption restricts its use to smooth estimators and makes it unreliable for nonsmooth statistics like the median or interquartile range. It also tends to underestimate variance for nonlinear statistics (cf [Shao and Tu \(1995\)](#); [Davison and Hinkley \(1997\)](#)).

The bootstrap, although more flexible, can be computationally intensive, especially when paired with complex models or large datasets (cf [Efron and Tibshirani \(1993\)](#); [Davison and Hinkley \(1997\)](#)). It can also perform poorly with small samples or highly dependent data unless modifications like the block bootstrap are applied (cf [Davison and Hinkley \(1997\)](#)).

Another key limitation is that both methods assume the observed sample adequately represents the population. If the original data are biased, contain outliers, or are not identically distributed, resampling can perpetuate those issues rather than correct for them.

6.5. *Conclusion on Practical Use.* In practice, the choice between jackknife and bootstrap depends on several factors, including sample size, estimator complexity, computational resources, and robustness requirements. Often, these methods are

used in combination: the jackknife for rapid, exploratory diagnostics and the bootstrap for final inference and reporting (cf [Efron and Tibshirani \(1993\)](#); [Shao and Tu \(1995\)](#); [Davison and Hinkley \(1997\)](#)).

7. Conclusion. This thesis presented an in-depth comparative analysis of the jackknife and bootstrap resampling methods, focusing on their theoretical foundations, methodological behavior, computational characteristics, and practical applications. By combining formal exposition with extensive simulation studies and application-based discussion, we have highlighted the respective strengths and limitations of both techniques.

The theoretical review showed that the jackknife offers a computationally efficient and deterministic approach suitable for smooth, linear estimators. Its leave-one-out design provides rapid diagnostics and reliable inference in simple settings. However, it lacks robustness for nonlinear or discontinuous estimators and may underestimate variance or bias in such cases (cf [Tukey \(1958\)](#); [Shao and Tu \(1995\)](#); [Davison and Hinkley \(1997\)](#)).

In contrast, the bootstrap, though more computationally demanding, proved to be far more flexible and broadly applicable. Its ability to simulate the empirical distribution of an estimator from resampled data enables accurate estimation of bias, variance, and confidence intervals—even for complex or unknown distributions (cf [Efron \(1979\)](#); [Efron and Tibshirani \(1993\)](#); [Davison and Hinkley \(1997\)](#)). The variety of interval construction techniques (e.g., percentile, basic, BCa) further enhances its utility (cf [Efron and Tibshirani \(1993\)](#); [Davison and Hinkley \(1997\)](#)).

The simulation study reinforced these findings. While the jackknife performed well for mean estimation under normal distributions, it struggled with skewed or heavy-tailed data and nonlinear statistics like the median. The bootstrap showed consistently strong performance across all scenarios, particularly for confidence interval coverage and bias correction.

From a computational perspective, the jackknife is preferable when speed and simplicity are critical, or when working with very large datasets or limited resources. The bootstrap, while slower, benefits from modern parallel computing environments, making it viable even for demanding tasks when high accuracy is essential.

Practical examples from diverse fields such as medicine, finance, ecology, and machine learning demonstrate the widespread relevance of both methods. Their adaptability makes them indispensable tools for statistical inference in the modern data-driven world.

In conclusion, both methods have distinct niches. The jackknife serves as a quick, efficient estimator diagnostic tool, while the bootstrap acts as a more robust, accurate inference engine. A thoughtful combination of both can often yield the best

balance of computational efficiency and statistical rigor. Future research may explore hybrid methods, such as jackknife-after-bootstrap, or extensions tailored to high-dimensional and dependent data settings, further expanding the potential of these powerful resampling techniques.

References.

- DAVISON, A. C. and HINKLEY, D. V. (1997). *Bootstrap Methods and Their Application*. Cambridge University Press, Cambridge.
- EFRON, B. (1979). Bootstrap Methods: Another Look at the Jackknife. *The Annals of Statistics* **7** 1–26.
- EFRON, B. (1982). *The Jackknife, the Bootstrap, and Other Resampling Plans*. SIAM, Philadelphia.
- EFRON, B. and TIBSHIRANI, R. J. (1993). *An Introduction to the Bootstrap*. Chapman & Hall/CRC, New York.
- HASTIE, T., TIBSHIRANI, R. and FRIEDMAN, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer, New York.
- HOROWITZ, J. L. (2001). *The Bootstrap* **5**. Elsevier, Amsterdam.
- PEDREGOSA, F., VAROQUAUX, G., GRAMFORT, A., MICHEL, V., THIRION, B., GRISEL, O., BLONDEL, M., PRETTENHOFER, P., WEISS, R., DUBOURG, V., VANDERPLAS, J., PASSOS, A., COURNAPEAU, D., BRUCHER, M., PERROT, M. and DUCHESNAY, E. (2011). Scikit-learn: Machine Learning in Python. <https://scikit-learn.org>. Accessed: 2025-05-04.
- QUENOUILLE, M. H. (1949). Approximate tests of correlation in time series. *Journal of the Royal Statistical Society. Series B (Methodological)* **11** 68–84.
- RAHMAN, F. (2019). EVALUATE PROBE SPEED DATA QUALITY TO IMPROVE TRANSPORTATION MODELING PhD thesis.
- SHAO, J. and TU, D. (1995). *The Jackknife and Bootstrap*. Springer, New York.
- R CORE TEAM (2025). R: A Language and Environment for Statistical Computing R Foundation for Statistical Computing, Vienna, Austria.
- SCIENCEDIRECT TOPICS Jackknife Resampling. <https://www.sciencedirect.com/topics/mathematics/jackknife-resampling>. Accessed: 2025-05-04.
- TUKEY, J. W. (1958). Bias and confidence in not quite large samples. *The Annals of Mathematical Statistics* **29** 614–623.
- WU, C. F. J. (1986). Jackknife, Bootstrap and Other Resampling Methods in Regression Analysis. *The Annals of Statistics* **14** 1261–1295.

Git: <https://github.com/JPK950/ResamplingMethods>

Bootstrap and Jackknife Estimations

```
set.seed(123)
data <- rnorm(30, mean = 5, sd = 2)

##### Part 3 #####

# Jackknife bias for mean
jack_means <- sapply(1:length(data), function(i) mean(data[-i]))
jack_means
jack_bias <- (length(data) - 1) * (mean(jack_means) - mean(data))
jack_bias

# Bootstrap bias for mean
B <- 1000
boot_means <- replicate(B, mean(sample(data, replace = TRUE)))
boot_means
boot_bias <- mean(boot_means) - mean(data)
boot_bias

n <- length(data)
jackknife_var <- (n - 1) / n * sum((jack_means - mean(jack_means))^2)
jackknife_var

boot_means <- replicate(1000, mean(sample(data, replace = TRUE)))
bootstrap_var <- var(boot_means)
bootstrap_var

hist(boot_means, col = "skyblue", main = "Bootstrap Distribution")
hist(jack_means, col = "pink", main = "Jackknife Distribution")

##### Part 4 #####

###4.2
# Simulate normal data
set.seed(123)
data <- rnorm(30, mean = 0, sd = 1)

# Bootstrap for mean
boot_means <- replicate(1000, mean(sample(data, replace = TRUE)))
boot_var <- var(boot_means)

# Jackknife for mean
```

```

jack_means <- sapply(1:30, function(i) mean(data[-i]))
jack_var <- (29/30) * sum((jack_means - mean(jack_means))^2)

mean(data)          # true sample mean
mean(boot_means)    # bootstrap estimate of the mean
boot_var            # bootstrap variance estimate
mean(jack_means)    # jackknife estimate of the mean
jack_var            # jackknife variance estimate

hist(boot_means, col = "skyblue", main = "Bootstrap Sample Means")

###4.3

# Simulate exponential data
set.seed(456)
data <- rexp(30, rate = 1)

# Bootstrap for median
boot_medians <- replicate(1000, median(sample(data, replace = TRUE)))
boot_var_median <- var(boot_medians)
boot_ci_median <- quantile(boot_medians, probs = c(0.025, 0.975))

# Jackknife for median
jack_medians <- sapply(1:30, function(i) median(data[-i]))
jack_var_median <- (29/30) * sum((jack_medians - mean(jack_medians))^2)
jack_bias_median <- (29) * (mean(jack_medians) - median(data))

# Print results
median(data)          # True sample median
mean(boot_medians)    # Bootstrap estimate of the median
boot_var_median       # Bootstrap variance
boot_ci_median        # Bootstrap 95% CI

mean(jack_medians)    # Jackknife estimate of the median
jack_var_median       # Jackknife variance
jack_bias_median      # Jackknife bias

### 4.4 Regression

# Simulate data
set.seed(789)
n <- 100
X <- runif(n, 0, 10)

# Homoscedastic case

```

```

epsilon_homo <- rnorm(n, mean = 0, sd = 1)
Y_homo <- 2 + 0.5 * X + epsilon_homo # True beta1 = 0.5

# Heteroscedastic case
epsilon_hetero <- rnorm(n, mean = 0, sd = sqrt(1 + 0.5 * X^2))
Y_hetero <- 2 + 0.5 * X + epsilon_hetero

# Function to estimate beta1 from linear regression
get_beta1 <- function(Y, X) {
  coef(lm(Y ~ X))[2]
}

# Jackknife estimate for beta1
jackknife_beta1 <- function(Y, X) {
  n <- length(Y)
  estimates <- sapply(1:n, function(i) get_beta1(Y[-i], X[-i]))
  beta_jack <- mean(estimates)
  var_jack <- (n - 1) / n * sum((estimates - beta_jack)^2)
  list(estimate = beta_jack, variance = var_jack)
}

# Bootstrap estimate for beta1
bootstrap_beta1 <- function(Y, X, B = 1000) {
  n <- length(Y)
  estimates <- replicate(B, {
    idx <- sample(1:n, replace = TRUE)
    get_beta1(Y[idx], X[idx])
  })
  list(estimate = mean(estimates),
       variance = var(estimates))
}

# Run for homoscedastic
jack_homo <- jackknife_beta1(Y_homo, X)
boot_homo <- bootstrap_beta1(Y_homo, X)

# Run for heteroscedastic
jack_hetero <- jackknife_beta1(Y_hetero, X)
boot_hetero <- bootstrap_beta1(Y_hetero, X)

# Output results
cat("Homoscedastic Model:\n")
cat("  Jackknife beta1:", round(jack_homo$estimate, 4),
    "| Variance:", round(jack_homo$variance, 5), "\n")
cat("  Bootstrap beta1:", round(boot_homo$estimate, 4),
    "| Variance:", round(boot_homo$variance, 5), "\n\n")

```

```
cat("Heteroscedastic Model:\n")
cat("  Jackknife beta1:", round(jack_hetero$estimate, 4),
    "| Variance:", round(jack_hetero$variance, 5), "\n")
cat("  Bootstrap beta1:", round(boot_hetero$estimate, 4),
    "| Variance:", round(boot_hetero$variance, 5), "\n")
```

RESPONSE TO REVIEWER COMMENTS

First of all, thank you very much for taking the time to read my thesis and for providing detailed comments that helped me improve the manuscript significantly. Below, I address each point individually.

Major Comments.

- **The Abstract before the Introduction is missing.**
The Abstract has now been added before the Introduction.
- **The given template is not used for this paper.**
The correct template is now used throughout the document.
- **The citations for the referenced literature are missing in the text.**
All citations have been added consistently in the text.
- **The 'References' heading appears twice.**
The duplicated 'References' heading has been removed.
- **The Appendix for figures and code is missing.**
The Appendix has been added, including all figures and code.
- **The Git-Repository with the code is missing.**
The Git repository link and information are included in the Appendix.
- **It would be helpful if Figure 1 on page 5 and Figure 2 on page 6 were explained in more detail rather than just inserted without commentary.**
The figures have now been explained in more detail and referenced explicitly in the text. However, unfortunately, the quality of Figures 1 and 2 could not be improved further, and the downloaded versions are also not optimal. If acceptable, I could redraw them myself for the final version.
- **In this paper, Table 1 and 2 could be referenced within the text, and the values could be briefly explained.**
Tables 1 and 2 are now referenced and discussed within the main text.
- **Table 3 could be briefly summarized in "Comparison Summary" and could also be referenced.**
Table 3 has been summarized in the "Comparison Summary" section and is now referenced appropriately.
- **A short explanation of Figure 3 and Figure 4 would be nice.**
Explanations of Figures 3 and 4 have been added in the corresponding sections.
- **In some instances, formulas are typed out, while in 3.3 and 3.4 they are inserted as images.**

All formulas have been unified to be typed consistently.

- **In section 4.4 a linear model is used and β_1 is estimated. How exactly were the data simulated and was OLS used for the estimation?**

A detailed explanation has been added describing how the data were simulated and confirming that OLS was used for estimation.

Minor Comments.

- **I think the section "Contents" is not needed in the paper.**

The "Contents" section has been removed.

- **It would be preferable if the figure descriptions cited the articles more precisely.**

The figure descriptions have been revised to cite the articles precisely.

- **It would make more sense if "Key Formulas for Bias and Variance Estimation" were before "Estimating Bias" and "Estimating Variance".**

The order of these sections has been adjusted accordingly.

- **Maybe the table number on page 11 (Table 4) could be included in the text.**

Table 4 is now explicitly referenced in the text.

- **In the paper are deleted jackknife, the jackknife after bootstrap, block bootstrap and wild bootstrap mentioned on the pages 5 and 6. Will they be explained in more detail?**

These methods have been described in more detail in the methodology section.

Typos.

- **On page 5, 'data set' is written, while in the rest of the paper 'dataset' is used.**

'dataset' is consistently used throughout the paper.

All changes have been implemented and new additions and explanations highlighted in the revised document for your convenience. Thank you again for your constructive feedback.