Comparing Methods for Constructing Confidence Intervals

A 95% Coverage Analysis of Four Approaches

Blessing Ofori-Atta

Introduction

Nonparametric bootstrap sampling provides a robust alternative to traditional parametric methods for statistical inference, particularly when the underlying distribution of the data is unknown or difficult to specify. Unlike parametric approaches, which rely on assumptions about the distributional form (e.g., normality), nonparametric bootstrap methods use the observed data itself to approximate the sampling distribution of a statistic. Despite these advantages, parametric methods, such as the Wald-based confidence interval, remain widely used in practice. When distributional assumptions are reasonably accurate and sample sizes are sufficiently large, parametric approaches can provide reliable results with greater computational efficiency. The Wald-based confidence interval, derived from inverting a large-sample Wald test, serves as a prime example of this. Efron (1993)[1] re-introduced the bootstrap principle, including the percentile and bootstrap-t methods. Hall (1987)[2] later proposed the symmetric bootstrap-t method, which modifies the bootstrap-t approach to produce symmetric confidence intervals. The percentile method chooses appropriate quantiles of an ordered bootstrap sample of coefficient estimates to form the interval, while the bootstrap-t method chooses appropriate quantiles of an ordered bootstrap sample of t-statistics to use as critical values in constructing the interval (Elias (2015)[3]). For more complex statistics without a simple standard error formula, a bootstrap estimate of the standard error can be computed for each bootstrap sample, requiring two levels of nested bootstrap sampling. The inner loop is used to estimate the standard error of the statistic of interest, while the outer loop is used to estimate the distribution of the statistic. This study compares variations of the symmetric bootstrap-t method—one using a plug-in estimate of the standard error and the other using a bootstrap estimate—with the percentile and Wald methods for constructing 95% confidence intervals for the population mean. The performance of these methods is evaluated through a simulation study across various sample sizes, using coverage rate as the metric. Coverage rate is the proportion of times the true parameter falls within the confidence interval, based on a specified number of replications, $R \gg 1$.

Description of the Solution Plan

A simulation study is conducted to explore the coverage rate of nominal 95% confidence intervals. The following methods are evaluated:

Symmetric Bootstrap-t (Plug-in Estimate of Standard Error)

1. Consistently estimate θ and $se(\widehat{\theta})$ using the observed sample (the simulated data): $\widehat{\theta}$, $\widehat{se}(\widehat{\theta})$.
2. Generate a bootstrap sample $X_1^\star, \ldots, X_n^*$.
3. Compute $\widehat{\theta}_b^\star$, $\widehat{se}_b^\star(\widehat{\theta})$, and

$$T_b^\star = \frac{\widehat{\theta}_b^\star - \widehat{\theta}}{\widehat{se}_b^\star(\widehat{\theta})}$$

   based on the bootstrap sample generated in step 1, where $\widehat{\theta}_b^\star$ and $\widehat{se}_b^\star(\widehat{\theta})$ are estimates of the parameter θ and its standard error.
4. Repeat steps 1 and 2 B times to obtain $T_b^\star, b = 1, \ldots, B$.
5. Order the bootstrap replications of T such that $|T_1^\star| \leq |T_2^\star| \leq \cdots \leq |T_B^\star|$. The absolute critical value, c, is determined as the $(1-\alpha)B$-th element. If αB is not an integer, c is the largest integer $<= (1-\alpha)(B+1)$. For $B = 999$ and $\alpha = 5\%$ this is the 950th ordered element. Now a $1 - \alpha$ confidence interval for θ is

$$\widehat{\theta} \pm |T_c^\star|.\widehat{se}(\widehat{\theta})$$

Symmetric Bootstrap-t (Boostrap Estimate of Standard Error)

1. Repeat steps 1-3. In step 3, instead of using the plug-in estimate of the standard error for the bootstrap sample, Generate another bootstrap sample from $X_1^\star, \ldots, X_n^*$ and estimate the standard error $\widehat{se}_b^\star(\widehat{\theta})$.
2. Repeats steps 4-5.

Percentile Method

1. Repeat steps 1-2. In step 2, compute $\widehat{\theta}_b^\star$ only.
2. Order the bootstrap replications of $\widehat{\theta}$ such that $\widehat{\theta}_1^\star \leq \widehat{\theta}_2^\star \leq \cdots \leq \widehat{\theta}_B^\star$. The lower and upper confidence bounds are the $B.\alpha/2$-th and $B(1 - \alpha/2)$-th ordered elements, respectively. A $1 - \alpha$ confidence interval for θ is $\left[\widehat{\theta}_{B.\alpha/2}^\star, \widehat{\theta}_{B(1-\alpha/2)}^\star\right]$.

Wald Method

1. Constructs a $1 - \alpha$ confidence interval for the population mean using the formula

$$\hat{\theta} \pm z_{1-\alpha/2}.\hat{se}(\hat{\theta})$$

where $z_{1-\alpha/2}$ is the $(1 - \alpha/2)$ quantile of the standard normal distribution.

Simulation Design

A simulation involved generating $n \in 10, 20, 50, 100, 500, 1000$ observations from an exponential distribution with mean one. For each sample size, $mcrep = 10,000$ replications of the observed data, were generated, and 95% confidence intervals were constructed using above methods. The coverage rate, defined as the proportion of replications in which the true mean ($\mu = 1$) fell within the confidence interval, was calculated.

Results

```
source("util.fun.R")

# set #sample sizes to us
n <- c(10, 20, 50, 100, 200)


# Function to simulate the mcreps of x for each n.
# Then, calculate the coverage of the confidence intervals for the 4 methods
simfun <- function(true.mean = 1, mcrep = 1000, nobs, alpha = 0.05){
  sim <- rexp(nobs * mcrep, rate = 1/true.mean) # Simulate data
  #sim <- rnorm(nobs * mcrep) # Simulate data
  xmat <- matrix(sim, ncol = mcrep) #each column is a replication
  xbar <- colMeans(xmat) # Compute the sample mean for each replication


  cl <- makeCluster(detectCores()-2)
  registerDoParallel(cl)
  sim_results <- foreach(i = 1:mcrep, .combine = rbind, .export = c("bootstrap", "theta", "c
    # Generate nboot independent bootstrap datasets and compute the 95% coverage
    boot.res <- bootstrap(x = xmat[, 1], xbar=xbar[1], theta=theta, nboot=1000)

    # Verify if the true mean is within the confidence interval
```

3

```r
    sim_cover.95 <- cover.95(
      x = xmat[, i],
      xbar = xbar[i],
      bootsamples = boot.res$bsample,
      thetastar = boot.res$thetastar,
      alpha = alpha,
      mu = true.mean)
  }

  # Stop the cluster
  stopCluster(cl)

  #Return the average coverage rate across all replications
  return(colMeans(sim_results))
}



# Repeat the simulation for each sample size
cl <- makeCluster(detectCores()-2)
registerDoParallel(cl)
coverage <- foreach(i = n, .combine = rbind, .packages = c("doParallel")) %dopar% {

  simfun(nobs = i)
}

# Stop the cluster
stopCluster(cl)
```
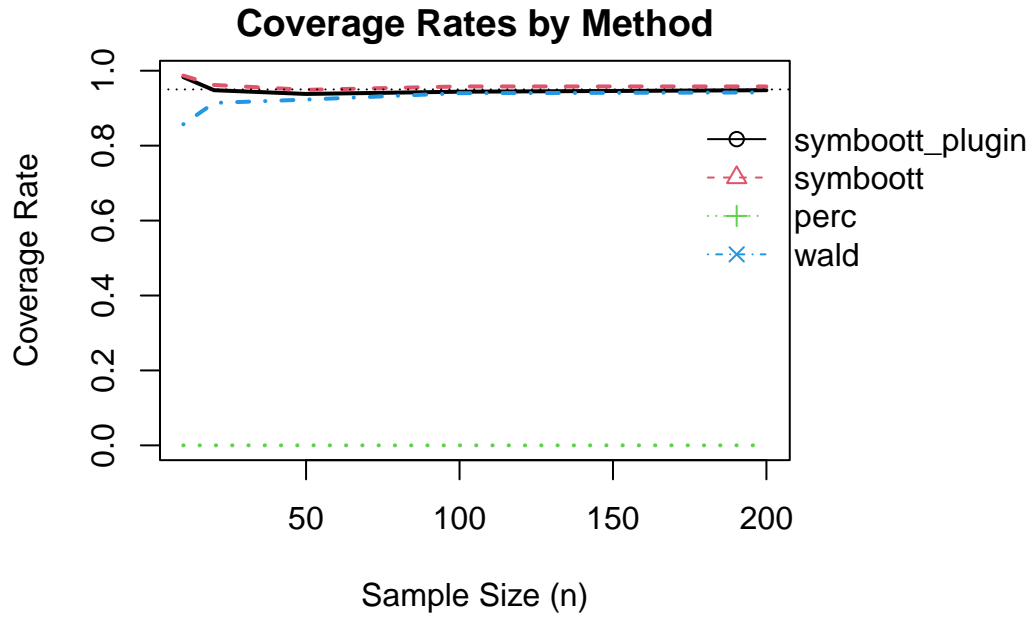
```r
coverage_matrix <- cbind(n, coverage)
# Plot the coverage rate
par(mar=c(5.1, 4.1, 2, 7))
matplot(coverage_matrix[, 1], coverage_matrix[, -1], type = "l", pch = 1:4, col = 1:4, lty =
        xlab = "Sample Size (n)", ylab = "Coverage Rate",
        main = "Coverage Rates by Method")

abline(h = 0.95, lty = 3)

par(xpd = TRUE)
legend("topright", inset = c(-0.4, 0.1), legend = colnames(coverage), col = 1:4, lty = 1:4,
```

Coverage Rates by Method

References

1. Efron, B., & Tibshirani, R. J. (1993). An introduction to the bootstrap. New York:Chapman & Hall.
2. Hall, P. (1988). On Symmetric Bootstrap Confidence Intervals. Journal of the Royal Statistical Society. Series B (Methodological), 50(1), 35–45. http://www.jstor.org/stable/2345806.
3. Elias, C. J. (2015). Percentile and percentile- t bootstrap confidence intervals: A practical comparison. Journal of Econometric Methods, 4(1), 153–161. doi:10.1515/jem-2013-0015
4. http://users.stat.umn.edu/~helwig/notes/npboot-notes.html