Mini Project 4

Yanjiao Jia (Contribution 50%)

Yuzhuang Feng (Contribution 50%)

Object

In this project, we build a function in R to compute nonparametric bootstrap confidence intervals for a population parameter based on a random sample from the population. We computed three bootstrap confidence intervals which are normal approximation, basic bootstrap, and percentile bootstrap. And we also computed bias and standard error of the parameter estimate which are needed for the confidence intervals.

Procedure and Analysis

1.Starting the computation, we need to construct our own bootstrap.

In our function, we made replacement and get the 90th percentile of every set of resample data and let it be resampled. And we made t0 be the 90th percentile of the original sample.

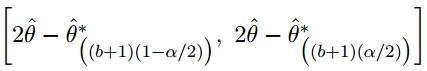
Then we used *bias<-mean(resample.t)-t0* to calculate the estimated bias and *sderror<-sd(resample.t)* to calculate the standard error of the 999 90th percentiles of every set of resample data for the later computation of confidence intervals.

And then we constructed

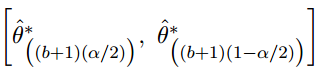
1. normal approximation CI



1. basic bootstrap CI



1. percentile bootstrap CI



We applied our own bootstrap function to the cpu data and printed all of the result computing from our own bootstrap.

2.The second step we did is using the boot package of R to do the calculation to verify the result we got before.

First we used boot strap function to get the 90th population percentile, estimated bias and standard error. And then we got three 95% bootstrap CI using boot.ci function and print out.

3.The result from computation using our own function is below:

bootstrap(cpu)

(1) parameter estimate: 82

(2) estimated bias: -3.242242

(3) standard error: 12.41634

(4) 95% normal approximation method: 60.90666 109.5778

(5) 95% percentile bootstrap method: 59 94

(6) 95% basic bootstrap method: 70 105

Comparing the result above with the result computed from boot package of R :

Bootstrap Statistics :

original bias std. error

t1\* 82 -3.790791 11.88308

Intervals :

Level Normal Basic

95% ( 62.50, 109.08 ) ( 70.00, 104.70 )

Level Percentile BCa

95% (59.3, 94.0 ) (59.0, 94.0 )

Calculations and Intervals on Original Scale

We can see two sets of result are very close. And three 95% bootstrap confidence intervals for the parameter are accurate. And the percentile bootstrap CI has the smallest width.

Bootstrap provide a good method to calculate the confidence interval when the population is not normal and the sample size is not large enough.

Code

# PART I

# Construct the our own bootstrap function to get parameter

# estimate (90th population percentile), estimated bias and

# standard error, the three 95% bootstrap confidence

# intervals for the parameter.

bootstrap<-function(x)

{

# get the 90th percentile of every set of resample data

resample.t<-replicate(999, quantile(sample(x, length(x), replace = TRUE, prob = NULL), .9, names = FALSE))

# get the 90th percentile of the original data

t0<-quantile(x, .9, names = FALSE)

# calculate the bias

bias<-mean(resample.t)-t0

# calculate the standard error of the 999 90th percentile of

# every set of resample data

sderror<-sd(resample.t)

# use 95% normal approximation method to calculate the bootstrap CI

normal<-c(t0-bias-qnorm(0.975)\*sderror,

t0-bias-qnorm(0.025)\*sderror)

# use 95% percentile bootstrap method to calculate the bootstrap CI

percentileBootstrap<-sort(resample.t)[c(25,975)]

# use 95% basic bootstrap method to calculate the bootstrap CI

basic<-c(2\*t0-percentileBootstrap[2], 2\*t0-percentileBootstrap[1])

# print out all the outcome

cat("(1) parameter estimate:",t0,"\n")

cat("(2) estimated bias:",bias,"\n")

cat("(3) standard error:",sderror,"\n")

cat("(4) 95% normal approximation method:",normal,"\n")

cat("(5) 95% percentile bootstrap method:",percentileBootstrap,"\n")

cat("(6) 95% basic bootstrap method:",basic,"\n")

}

library(boot)

cpu<-scan(file="cputime.txt")

bootstrap(cpu)

# PART II

# Then we use the boot package of R to do the calculation to verify our outcomes

# set the function to get 90th percentile

P <-function(x,indices)

{

result<-quantile(x[indices], .9,numes=FALSE)

return(result)

}

# use boot function to get the parameter estimate (90th

# population percentile), estimated bias and standard error

P.boot <-boot(cpu, P, R=999, sim="ordinary", stype="i")

print(P.boot)

# use boot.ci function to get the three 95% bootstrap

# confidenceintervals

print(boot.ci(P.boot))