# Simulation

par(mfrow=c(2,3))

grid <- seq(-3,3, length=1000)

cdf <- pnorm(grid)

plot(grid, cdf, type="l",xlab = "x", ylab = "cdf", sub="True cdf")

num <- 100

x <- rnorm(n)

x <- sort(x)

cdf.hat <- (1:num)/num

plot(x,cdf.hat,type="s",xlab="x",ylab="cdf", sub="emperical cdf", xlim=c(-3,2))

plot(grid, cdf, type = "l", xlab = "x", ylab = "cdf")

lines(x,cdf.hat,lty=3, col=3, lwd=3, type = "s")

alpha <- .05

eps <- sqrt(log(2/alpha)/(2\*num))

l <- pmax(cdf.hat - eps, 0)

u <- pmin(cdf.hat + eps, 1)

plot(grid, cdf, type = "l", xlab = "x", ylab="cdf")

lines(x,l,lty=2,col=2,type="s")

lines(x,u,lty=2,col=2,type="s")

library(boot)

foo = c(8,10,7,12,13,8,10,50)

my.mean = function(x,indices) {

return( mean(x[indices]) )

}

boot.out = boot(foo,my.mean,1000)

boot.ci(boot.out)

data <- c(8,10,7,12,13,8,10,50)

mean(data)

library(boot)

my.mean = function(x,indices) {

return( mean(x[indices]) )

}

boot.out = boot(foo,my.mean,1000)

boot.ci(boot.out)

75+2\*sd(boot.out$t)

c(mean(data)-1.96\*4.756461,mean(data)+1.96\*4.756461)

quantile(boot.out$t,0.025)

quantile(boot.out$t,0.975)

library(ggplot2)

ggplot(data.frame(x = boot.out$t), aes(x = x)) + geom\_density()+

geom\_vline(xintercept=c( 5.41, 24.07),color="green")+

geom\_vline(xintercept=c( 4.25, 20.88),color="blue")+

geom\_vline(xintercept=c(8.62, 25.25),color="purple")+

geom\_vline(xintercept=c( 9.12, 30.50 ),color="yellow")

# The BCa confidence interval is much high than the other three confidence intervals. It seems

# that this may be so since the BCa interval adjusts for bias and skew. Our distribution has these

# features, and thus, it seems natural that the BCa interval would be different than the other intervals.

# Application

x<-quakes$mag

par(mfrow=c(2,1))

hist(x,main="Earthquake Fiji data")

u<-ecdf(x) # this is a function

plot(u)

alpha <- 0.05

n <- length(x)

epsn<-sqrt(log(2/alpha)/(2\*n))

r<-max(x)-min(x)

grid<-seq(from=min(x)-0.01\*r,to=max(x)+0.01\*r,l=1000)

low.cdf<-pmax(u(grid)-epsn,0)

up.cdf<-pmin(u(grid)+epsn,1)

lines(grid,low.cdf,col="red")

lines(grid,up.cdf,col="red")

# suppose th = F(4.9)-F(4.3), then we can estimate th by th.hat as follows

a<-4.3

b<-4.9

print(th.hat<-u(b)-u(a))

# a confidence interval for th can be obtained from the binconf function

# within the library Hmisc

library(Hmisc)

tot<-sum( (x<=4.9) & (x>4.3))

binconf(tot,length(x),method="wilson",alpha)