Math 159, Spring 2010
Jo Hardin
R code for HW & final

Hesketh and Everitt (2000) report on a study by Caplehorn and Bell (1991) that investigated the times that heroin addicts remained in a clini for methadone maintenance treatment. The data in `heroin.txt` include the amount of time that the subjects stayed in the facility until treatment was terminated (column 4). For about 37% of the subjects, the study ended while they were still the in clinic (status=0). Thus, their survival time has been truncated. For this reason we might not want to estimate the mean survival time, but rather some other measure of typical survival time. Below we explore using the median as well as the 25% trimmed mean.

```
obs.stat<-median(heroin[,4]) # 367.5
obs.stat2<-mean(heroin[,4],trim=.25) # 378.3

heroin.rs<-sample(heroin[,4],238,replace=T)

median(heroin.rs)
mean(heroin.rs,trim=.25)
```

- Bootstrapping

```
test.stat<-c()
test.stat2<-c()
sd.test.stat<-c()
sd.test.stat2<-c()

for (i in 1:1000){
heroin.rs<-sample(heroin[,4],238,replace=T)
test.stat<-c(test.stat,median(heroin.rs))
test.stat2<-c(test.stat2,mean(heroin.rs,trim=.25))

test.stat.rs<-c()
test.stat2.rs<-c()

for (j in 1:1000){
heroin.rsrs<-sample(heroin.rs,238,replace=T)
test.stat.rs<-c(test.stat.rs,median(heroin.rsrs))
        test.stat2.rs<-c(test.stat2.rs,mean(heroin.rsrs,trim=.25))
}
sd.test.stat<-c(sd.test.stat,sd(test.stat.rs))
sd.test.stat2<-c(sd.test.stat2,sd(test.stat2.rs))
}

par(mfcol=c(2,3))
hist(heroin[,4])
hist(heroin.rs)
hist(test.stat)
qqnorm(test.stat)
qqline(test.stat)
hist(test.stat2)
qqnorm(test.stat2)
qqline(test.stat2)

mean(test.stat) # 374.522
sd(test.stat) # 33.281

mean(test.stat2) # 378.040
sd(test.stat2) # 23.407
```

OR

```
bs.proc <- bootstrap(heroin[,4],1000,median)
bs.proc2 <- bootstrap(heroin[,4],1000,mean,trim=.25)
```

- 95% Percentile CI

```
c(sort(test.stat)[25],sort(test.stat)[975])
c(sort(test.stat2)[25],sort(test.stat2)[975])
```

OR

```
c(sort(bs.proc$thetastar)[25],sort(bs.proc$thetastar)[975])
c(sort(bs.proc2$thetastar)[25],sort(bs.proc2$thetastar)[975])
```

- 95% t-CI

```
c(obs.stat - qt(.975,237)*sd(test.stat), obs.stat + qt(.975,237)*sd(test.stat))
c(obs.stat2 - qt(.975,237)*sd(test.stat2), obs.stat2 + qt(.975,237)*sd(test.stat2))
```

OR

```
se.bs <- sd(bs.proc$thetastar)
se.bs2 <- sd(bs.proc2$thetastar)

c(obs.stat - qt(.975,237)*se.bs, obs.stat + qt(.975,237)*se.bs)
c(obs.stat2 - qt(.975,237)*se.bs2, obs.stat2 + qt(.975,237)*se.bs2)
```

- 95% Bootstrap-t CI

```
t.hat<-(test.stat - obs.stat)/sd.test.stat
t.hat2<-(test.stat2 - obs.stat2)/sd.test.stat2

c(obs.stat - sort(t.hat)[975]*sd(test.stat), obs.stat - sort(t.hat)[25]*sd(test.stat))
c(obs.stat2 - sort(t.hat2)[975]*sd(test.stat2), obs.stat2 - sort(t.hat2)[25]*sd(test.stat2))
```

OR

```
boott(heroin[,4],1000,median)
boott(heroin[,4],1000,mean,trim=.25)
```

- 95% BCa interval

```
test.stat.jk<-c()
test.stat2.jk<-c()

for(i in 1:length(heroin[,4])){

test.stat.jk<-c(test.stat.jk,median(heroin[-i,4]))
 test.stat2.jk<-c(test.stat2.jk,mean(heroin[-i,4],trim=.25))
}

zo.hat<-qnorm(sum(test.stat<obs.stat)/1000,0,1)
a.hat<- sum((mean(test.stat.jk) - test.stat.jk)^3)/
                 (6*(sum((mean(test.stat.jk)-test.stat.jk)^2)^1.5))

zo.hat2<- qnorm(sum(test.stat2< obs.stat2)/1000,0,1)
a.hat2<- sum((mean(test.stat2.jk) - test.stat2.jk)^3)/
                 (6*(sum((mean(test.stat2.jk)-test.stat2.jk)^2)^1.5))

alpha1.bca<-pnorm(zo.hat + (zo.hat + qnorm(.975))/(1 - a.hat*(zo.hat + qnorm(.975))))
alpha2.bca<-pnorm(zo.hat + (zo.hat + qnorm(.025))/(1 - a.hat*(zo.hat + qnorm(.025))))


alpha1.bca2<-pnorm(zo.hat2 + (zo.hat2 + qnorm(.975))/(1 - a.hat2*(zo.hat2 + qnorm(.975))))
alpha2.bca2<-pnorm(zo.hat2 + (zo.hat2 + qnorm(.025))/(1 - a.hat2*(zo.hat2 + qnorm(.025))))


c(sort(test.stat)[ceiling(1000*alpha2.bca)],sort(test.stat)[ceiling(1000*alpha1.bca)])
c(sort(test.stat2)[ceiling(1000*alpha2.bca2)],sort(test.stat2)[ceiling(1000*alpha1.bca2)])
```

OR

```
bcanon(heroin[,4],1000,median)
bcanon(heroin[,4],1000,mean,trim=.25)
```