MA152_MW9f20.Rmd

Ethan Ashby

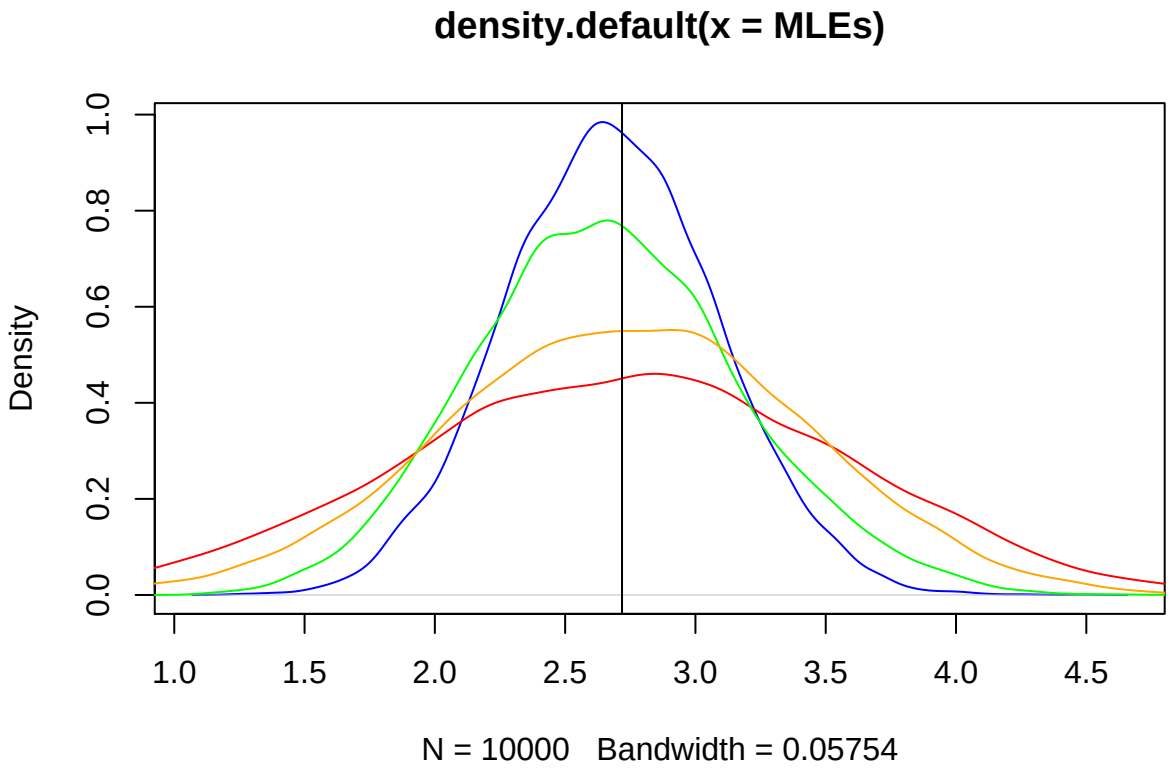
10/22/2020

b. Use a simulation to compare the four estimators above with respect to bias, variance, and MSE. Answer the following questions in your comparison:

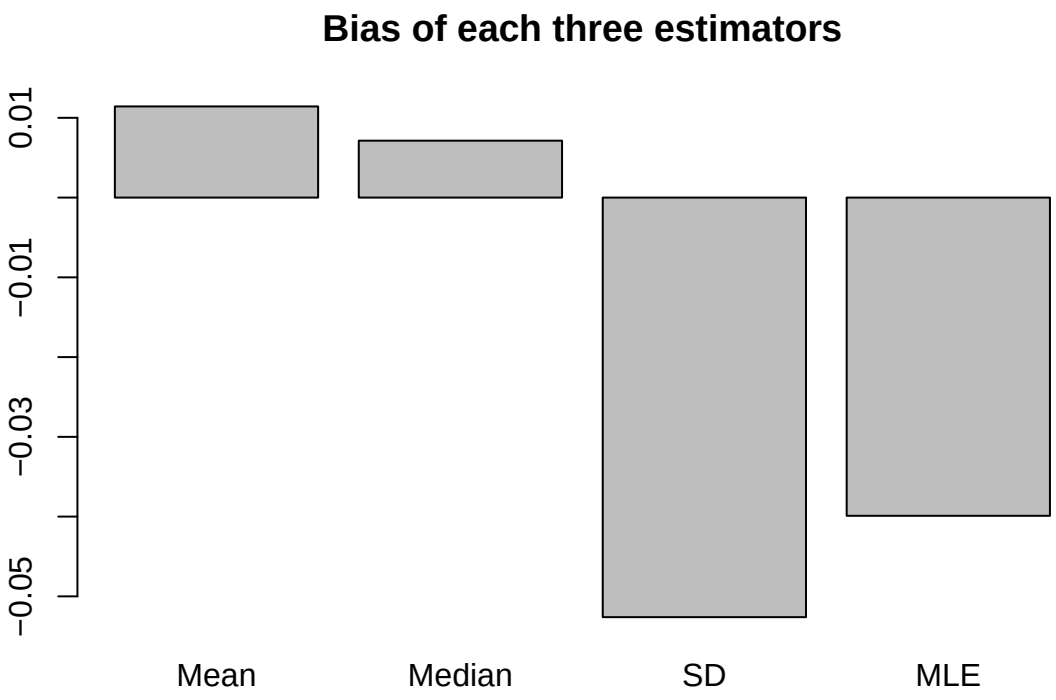
   i. Which estimator is (empirically) least biased? The median is empirically the least biased estimator per the plot shown below.
   ii. Which estimator has lowest empirical variability? Do any of the estimators reach the CRLB (assume unbiasedness)? The MLE, by a good margin, has the lowest empirical variability, which matches the CRLB.
   iii. Which estimator has lowest empirical MSE? The MLE has the lowest empirical MSE
   iv. Are you comfortable with the idea of using a normal distribution to describe the sampling distribution for any/all of the estimators? Explain. I'm fine with using normal approximations for the mean (b/c means are normally distributed), the median (b/c in the underlying data distribution the mean=median). I'm less comfortable with the other two being normal too, but per the density plots, it appears every estimator has a relatively bell-shaped sampling distribution.
   v. Which estimator would you recommend for the given setting? Explain. I'd suggest the MLE. It is not super biased and it's variance is the smallest theoretical value it could be. It also has the lowest empirical MSE, and appears to conform to a normal model, which could make inference pretty nice!

```
# Use sample size n = 15. keep this set-up in the first 3 lines
set.seed(11)
n.obs <- 15
n.samps <- 10000
theta <- exp(1)


means <- numeric(n.samps)
medians <- numeric(n.samps)
sds <- numeric(n.samps)
MLEs <- numeric(n.samps)


for (i in 1:n.samps){
  dat<-rnorm(n.obs, theta, sd=theta)
  means[i]<-mean(dat)
  medians[i]<-median(dat)
  sds[i]<-sd(dat)
  MLEs[i]<- -mean(dat)/2 + sqrt(sum(dat^2)/n.obs + (mean(dat)^2/4))
}

#Distributions of estimators
plot(density(MLEs), col="blue")
```

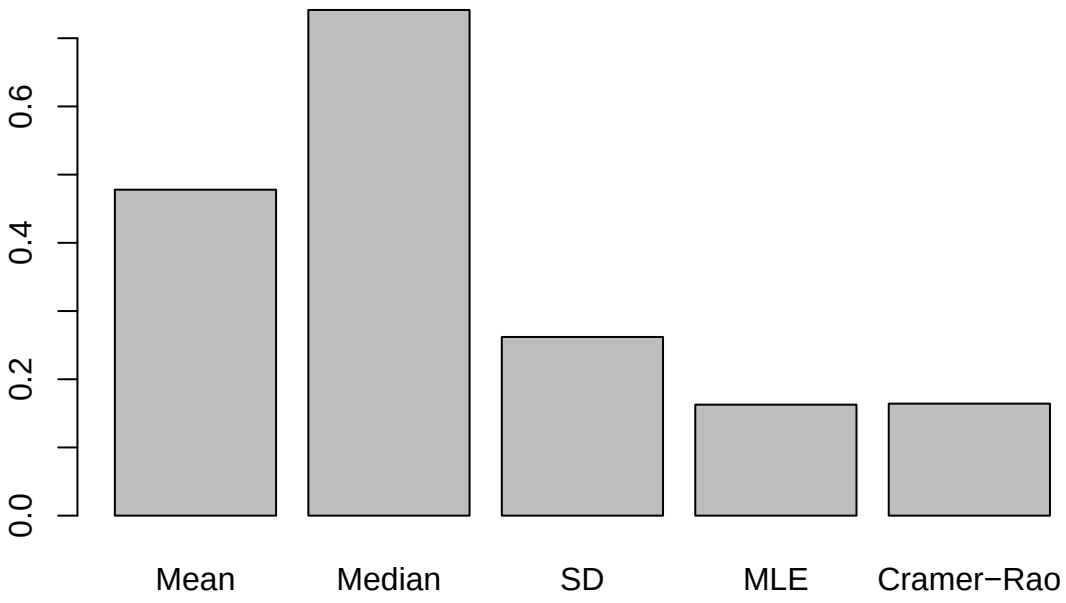
```
lines(density(medians), col="red")
lines(density(sds), col="green")
lines(density(means), col="orange")
abline(v=theta)
```




```
#bias
barplot(c(sum(means-theta)/n.samps,
sum(medians-theta)/n.samps,
sum(sds-theta)/n.samps,
sum(MLEs-theta)/n.samps), names= c("Mean", "Median", "SD", "MLE"), main="Bias of each three estimators")
```




```
#variability
barplot(c(var(means),
var(medians),
var(sds),
var(MLEs),
theta^2/(3*n.obs)), names= c("Mean", "Median", "SD", "MLE", "Cramer-Rao"), main="Variance of each three
```




```
#MSE
barplot(c(var(means)+(sum(means-theta)/n.samps)^2,
```

```r
var(medians)+(sum(medians-theta)/n.samps)^2,
var(sds)+(sum(sds-theta)/n.samps)^2,
var(MLEs)+(sum(MLEs-theta)/n.samps)^2), names= c("Mean", "Median", "SD", "MLE"), main="MSE of each three
```