Math 152 - Statistical Theory - Homework 7

write your name here

Due: 11/2/2018

Book problems:

8.8 – 4, 5 (AND show that the unbiased estimator of σ^2 is not efficient), 14 (you don't actually need to find the MLE), 16
8.9 – 15, 18

R problem

We are going to study almost, but not exactly, the same model as in class. The model for this problem is normal with mean θ and variance θ^2 (not θ as in the example in the notes).

The results from class about properties of MLEs are asymptotic. What happens in small samples?

The estimators of θ we wish to compare are:

- the sample median
- the sample mean
- the sample standard deviation times the sign of the sample mean
- the MLE

1. Turns out that the Fisher Information in θ is $3/\theta^2$. Show (with pencil) that the MLE of θ is

$$\hat{\theta} = -\overline{x}/2 + \sqrt{\left(\sum x_i^2\right)/n + \overline{x}^2/4}$$

.

2. Use a simulation to compare the four estimators above with respect to bias, variance, and MSE. Answer the following questions in your comparison:
   (a) Which estimator is (empirically) least biased?
   (b) Which estimator has lowest empirical variability? Do any of the estimators reach the CRLB (assume unbiasedness)?
   (c) Which estimator has lowest empirical MSE?
   (d) Are you comfortable with the idea of using a normal distribution to describe the sampling distribution for any/all of the estimators?
   (e) Which estimator would you recommend for the given setting?

```
# Use sample size n = 15.
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){
  # generate some data
  # means[i] <- mean(the data you generated)
  # etc.
}
```

You can write alternative code for calculating relevant characteristics of the distribution and displaying it, but I choose to put it together in a tidy framework like this:

```
library(tidyverse)

est.info <- data.frame(value = c(means , ...),
                       type = c(rep("mean", n.samps), ...) )

est.info %>%
  group_by(type) %>%
  summarize(est.bias = mean(value) - exp(1), est.mean = mean(value),
            est.var = var(value), est.sd = sd(value)) %>%
  mutate(est.mse = est.var + est.bias^2)


est.info %>%
  ggplot(aes(x=value, color = type)) + geom_density() +
  geom_vline(xintercept = exp(1))
```