

Math 152 - Statistical Theory - Homework 5

write your name here

Due: 10/5/2018

Book problems:

7.6 – 5, 6, 8, 9, 19, 21, 23

R problem (Tanks)

Consider the tank problem encountered in class. Your task at hand is to provide the best possible estimator for the true number of tanks. Consider the R code below that analyzes two estimators ($2\bar{X}$ and the maximum of the sample). You should **provide an argument (in words but using the evidence collected here) for your choice of estimator** using the following information:

- a comparison of at least 5 estimators
- consideration of sample bias, sample variance, sample median, sample mean, sample MSE
- run the analysis twice. Once with a sample size of $n=5$, once with a sample size of $n=100$

```
# Keep the population size at 447
npop = 447
nsamp = 5 # change this to 100 for the second part of the analysis
reps = 10000

xbar2 = c() # placeholder for your repeated sample statistics
sampmax = c()

for (i in 1:reps){
  mysample = sample(1:npop,nsamp,replace=F) # sample some tanks from the population
  xbar2 = c(xbar2, 2*mean(mysample) )
  sampmax = c(sampmax, max(mysample))
}

estimate <- c(xbar2, sampmax)
method <- c(rep("2xbars", reps), rep("samplemax", reps))
all.estimates <- data.frame(estimate, method)

# below is some syntax with which you may be unfamiliar.
# if you don't have the packages, you may need to run, for example:
# install.packages("dplyr")

# also, if you create the statistics as above, all of the code below should
# work nicely without you doing anything to it.

# please ask if you have question!

library(dplyr)
library(ggplot2)

all.estimates %>%
```

```
group_by(method) %>%
  summarize(mean = mean(estimate), median = median(estimate), bias = mean(estimate) - npop,
            var = var(estimate), mse = (mean(estimate) - npop)^2 + var(estimate))
```

```
## # A tibble: 2 x 6
##   method    mean median  bias   var   mse
##   <fct>    <dbl> <dbl> <dbl> <dbl> <dbl>
## 1 2xbars    448.  447.   1.37 13310. 13312.
## 2 samplemax 374.  391 -73.0  3913.  9243.
```

```
ggplot(all.estimates, aes(x = estimate)) + geom_histogram() +
  geom_vline(xintercept = npop) + facet_wrap(~method)
```

