2020 Summer STAT 220 Final Exam, Part 3

Your Name Goes Here

You can use this R markdown file as a tempelate for your work, though you can also create your own RMD file. If you have trouble knitting it to pdf directly, you can first knitting it to html or Word and then convert it into pdf. You can also type your answer in a Word document and copy and paste the R codes and output/plots into it.

The questions start below.

We will work on the North Carolina birth record data in Lab 8 http://www.stat.uchicago.edu/~yibi/s220/labs/lab08.html that we have looked at several times in assignments. It is a random sample of 1000 birth records of all babies born in the state of North Carolina in 2004. You can load the data set by running the command below in R directly without changing the working directory. Please make sure you have internet connection before running the command below.

```
nc = read.csv("https://www.openintro.org/stat/data/csv/ncbirths.csv")
```

We are interested in comparing the average ages of married mothers and unmarried mothers that gave birth in the state of North Carolina in 2004. In the data, the variable `mage` stores the age of the mother in years, and the variable `marital` indicator whether the mother is married or unmarried.

Q1 — 3 Points

Using R, find out the following summary statistics.

- the number of cases that the mother is married.
- the number of cases that the mother is unmarried.
- the averae age of the married mothers in the data
- the averae age of the unmarried mothers in the data
- the standard deviation (SD) of the age of the married mothers in the data
- the standard deviation (SD) of the age of the married mothers in the data

```
# Please show the R command(s) used here
```

Q2 — 4 Points

Please construct a 95% confidence interval for the difference in the means of the mother's age between married mothers and unmarried mothers who gave birth in the state of North Carolina in 2004. Show your work. If you use R, please show your R codes.

Q3 — 3 Points

Please check the necessary condition(s) required to construct the confidence interval in the previous part.

```
# Please show the R command(s) used here
```