STAT 630, HW 4
Due: Thursday, September 23

Reading: *OpenIntro*, Section 4.1. For a review of random variables, expectation, and variance read Section 3.4.

Directions: Please submit your completed assignment to Blackboard. For the concept questions, your solutions may be typed, or handwritten and then scanned. The data analysis questions should be completed using R Markdown and then rendered to HTML, PDF, or Word format.

Concept Questions

Please refer to the lecture 4 notes.

Exercise 1. Heights of 10 year olds, regardless of gender, closely follow a normal distribution with mean 55 inches and standard deviation 6 inches.

(a) What is the probability that a randomly chosen 10 year old is shorter than 48 inches?

(b) What is the probability that a randomly chosen 10 year old is between 60 and 65 inches?

(c) What is the height cutoff for the tallest 5% of ten year olds (the 95^{th} percentile)?

Exercise 2. Josh, a student in a statistics class, scored 85 points on the first exam and 80 points on the second exam. The mean score on the first exam for all students in this course was 79 with a standard deviation of 4. The mean score on the second exam was 60 with a standard deviation of 10. The distributions of both exam scores are approximately normal.

(a) What is Josh's z-score on the first exam?

(b) What is Josh's z-score on the second exam?

(c) On which exam did Josh do better when compared with other students in the class? Explain.

Exercise 3. Suppose weights of the checked baggage of airline passengers follow a nearly normal distribution with mean 45 pounds and standard deviation 3.2 pounds. Most airlines charge a fee for baggage that weigh in excess of 50 pounds. Determine what percent of airline passengers incur this fee.

Exercise 4. Assume a random variable X follows a normal distribution with mean μ and standard deviation σ. What is the probability that an observation falls below $\mu + 2\sigma$?

Data Analysis and R Questions

Please refer to lab 4.

Exercise 5. Let $Z \sim N(0,1)$ be a random variable following a standard normal distribution. Use `qnorm()` to find the value c such that:

(a) $P(Z < c) = 0.8$

(b) $P(Z > c) = 0.01$

(c) $P(-c < Z < c) = 0.95$

Exercise 6. Generate 1000 random numbers from a normal distribution with mean $\mu = 100$ and standard deviation $\sigma = 20$, and then plot the histogram. Also make a normal QQ plot with the random numbers you generated.

Exercise 7. Use `dnorm()` to plot normal density curves for the distributions $N(\mu = 20, \sigma = 2)$ and $N(\mu = 20, \sigma = 4)$. Include both density curves on the same plot. Comment on how changing the standard deviation parameter σ affects the shape of the normal curve.

The following exercise uses the CDC data set. Run the following command to load this data set into R:

```
cdc <- readRDS(url("https://ericwfox.github.io/data/cdc.rds"))
```

Exercise 8. Subset the females from the `cdc` data set. Make a density histogram and normal QQ plot of the `height` and `weight` variables for the females. Superimpose a normal curve on the density histogram, using the sample mean and sample standard deviation as the parameters. Use `qqline()` to add a reference line to the QQ plot. Do the points on the QQ plots fall on the straight line? Comment on any deviations in the data from the normal distribution.