STA 100 Homework 2

Due 11:59 pm Friday, July 14 onto Gradescope

1. For a standard normal random variable $Z \sim N(0, 1)$, find the following:
   (a) $P(Z < 1.2)$.
   (b) $P(Z \geq -2.3)$.
   (c) $P(-1 < Z \leq 0.31)$.
   (d) The 97.5th percentile of Z.

2. Assume that heart rate (in beats per minute, or bpm) before an exam for STA 100 students is distributed normal, with a mean of 95 bpm and a standard deviation of 18.5 bpm. Assume all students in the following problem are selected from this population.
   (a) Find the probability that the heart rate of a randomly selected student is above 110.
   (b) What is the probability that a randomly selected student has a heart rate between 90 and 120?
   (c) What is the first quartile of heart rates for randomly selected students?
   (d) What is the third quartile of heart rates for randomly selected students?
   (e) What is the 8th percentile for heart rates among randomly selected students?
   (f) If we know a randomly selected students heart rate is over 100 (it is given), what is the probability that it is under 125? Hint: Conditional probability.

3. Mensa is an organization that allows people to join only if their Stanford-Binet IQs are in the top 2% of the population. Assume the population mean of Stanford-Binet IQs is 100, and the variance is 225, and that the population is normally distributed.
   (a) What is the lowest Stanford-Binet IQ you could have and still be eligible to join Mensa?
   (b) What is 99th percentile for the Stanford-Binet IQ scores?
   (c) What is the probability that the Stanford-Binet IQ of a randomly selected person is between 85 and 115?
   (d) What is the probability that two randomly selected people both have Stanford-Binet IQs which qualify them for Mensa?

4. You must use properties of linear combinations of random variables to solve these problems.
   (a) Let X be a random variable where $\mu_X = 23$. Find the mean of Z, where $Z = -2 + 4X$.
   (b) Let X and Y be two random variables where $\mu_X = 23, \mu_Y = 6$. Find the mean of Z, where $Z = X + 2Y$.
   (c) Let X be a random variable where $\sigma_X^2 = 1/4$. Find the variance of Z, where $Z = -10 - 2X$.
   (d) Let X and Y be two *independent* random variables where $\sigma_X^2 = 2, \sigma_Y^2 = 1/4$. Find the variance of Z, where $Z = X - 2Y$.

5. The mean daily rainfall between January 1, 2010, through January 1, 2012, at Sunshine City, was 0.01 inches with a standard deviation of 0.1 inches. Based on this information, do you think it is reasonable to believe that daily rainfall at Sunshine City follows a normal distribution? Explain. (Hint: Think about the possible values for daily rainfall and use the 68%-95%-99.7% rule.)

R is necessary for the remaining questions. Attach source codes and any plots you produce to your homework submission. You may write down your numerical results.

6. (a) Consider a binomial random variable $X \sim B(1000, 0.5)$, calculate $P(X \leq 489)$ and $P(X > 510)$. Hint: `pbinom()`.

   (b) Generate a random sample of size $n = 500$ from the binomial distribution $B(1000, 0.5)$ using the following command.

   ```
   rbinom(500, size = 1000, prob = 0.5)
   ```

   Use `quantile()` to find the first and third quartiles of this sample.

   (c) Create a normal quantile plot for this sample using `stat_qq()` and `stat_qq_line()` from `ggplot2` package. Does the sample agree with the normal distribution? Hint: See Page 7 of the handout Ch4.pdf.

7. The dataset we will be exploring is the `vitamina` data in `isdals` package, The daily food intake was studied for 2224 subjects, and the content of many different vitamins and substances were measured. You can use the following commands to load the `vitamina` data into your working environment.

   ```
   library(isdals); data(vitamina)
   ```

   An alternative way is to download the data set from Canvas (Files → Data → vitamina.csv) and import it into R using `read.csv()` function.

   (a) Create a normal quantile plot for the variable `wt`, which represents the weight of each subject in kilograms. Based on the sample, does it appear to follow a normal distribution? If not, what type of skewness does the data exhibit?

   (b) Apply a logarithmic transformation to the variable `wt`. Create a normal quantile plot for the transformed sample. Compare this plot with the normal quantile plot in (a). What conclusion can you draw based on the comparison?