STAT651 Homework 3

This HW reviews the binomial distribution, normal distribution, validity of parameters based on probabilities and QQplots.

(1) 10 years ago, public opinion on same-sex marriage was divided $50:50$.

A recent pole of 1000 randomly sampled (using an SRS) individuals was taken. In that poll, 560 of the 1000 said they supported same sex marriage.

(a) Suppose that public opinon on same-sex marriage has not changed and is still divided 50:50 (this means the poportion of adults who support same sex marriage is 50%). Calculate the chance that in a sample of 1000 **560 people or more** say they support same sex marriage.

Hint: Use a Binomial distribution and appropriate software.

(b) Based on your answer in part (a), what do you think about the underlying assumption that public opinion remains divided 50:50?

(c) A Sociologist tells you that he believes that public opinion has changed to 55%. Under the assumption that 55% of the general public support same-sex marriage, calculate the chance in a sample of 1000 **560 people or more** say they support same sex marriage.

(d) Using your answer in (c) is the Sociologist's claim viable? Does the data prove that public opinion is now 55%?

(2) To this question you will need to follow the steps in Lecture 8, slide 14-18.

Using JMP obtain the binomial probabilities when $p = 0.15$ and $n = 10, 50$ and 200. Remember for each n you will need a column with $\{0, 1, \ldots, n\}$, these are all the possible outcomes for the sum S_n.

For each of the S_n values ($x \in \{0, 1, \ldots, 10\}$, $S_n = \{0, 1, \ldots, 50\}$ and $x \in \{0, 1, \ldots, 200\}$) make the z-transform

$$z = \frac{x - (n \times 0.15)}{\sqrt{n \times 0.15 \times 0.85}},$$

(exactly what is done in slide 17). Doing this in JMP can be fiddly. This transforms all the different sums to the same range (between about -10 to 10).

Make a bar chart for each plot by going to Graph > Graph Builder. Drag and Pull Z to the x-axis and the Press the bar button. An example is given below in figure 1

- Make plots similar to that in Figure 1 for $n = 10, 50$ and 200. Take screen shots of each plot and put them next to each other. Describe what happens to the distribution as n gets larger.

(3) Suppose female heights are normally distributed with mean $\mu = 64.5$ and standard deviation 3 inches.

(a) Calculate the proportion of females more than 63 inches?

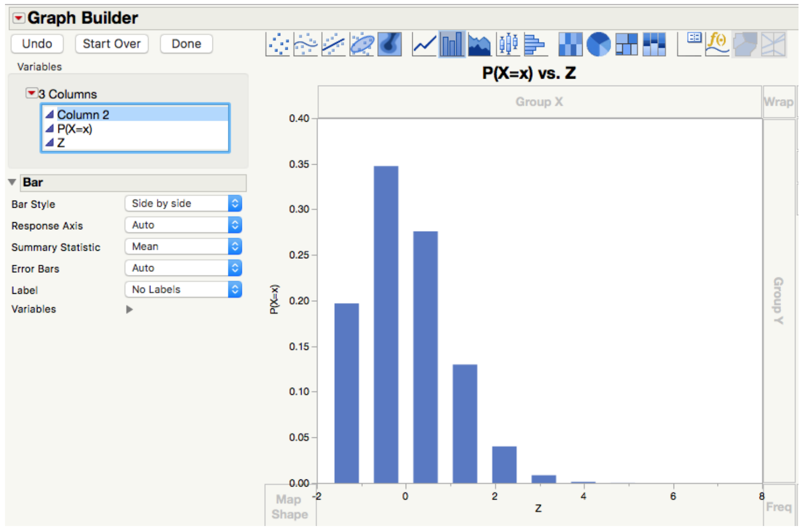

Figure 1: Making a plot using Graph Builder

(b) Construct an interval **centered** about the mean which contains 85% of female heights.

(c) Suppose a female is in the 90th percentile. How tall is she?

(4) If a machine is packing tomatoes correctly, then the average weight of a pack of tomatoes in a batch follows a normal distribution with mean 227grams and standard deviation 4grams.

A quality control agent checks the average weight in each batch packed. The average weight of one batch is 240g.

(a) Calculate the chance the average weight is 240grams or over, when the machine is packing tomatoes correctly.

(b) The chance calculated in part (a) should be quite small. Based on this probability, do you think the machine is functioning correctly?

(5) Input the data

`https://www.stat.tamu.edu/~suhasini/teaching651/old_faithfuldat.txt` into JMP or the statistical software you use.

(a) Make a QQplot of the eruption times. Describe the QQplot and how close it is to normality.

(b) Suppose you used the normal distribution to calculate the proportion of eruption times within a certain time interval. Based on the QQplot in (a) will the calculations be close to the true probabilities?