**Michael Rimmey**

**Professor Jeffrey Ziegler**

**2 April 2019**

**Quantitative Political Methods**

**Problem Set 3**

**Question 1**

R Code**:**

1. ObamaPerHand = newhamp[newhamp$pObama & newhamp$votesys=="H",]
3. ObamaPerMach = newhamp[newhamp$pObama & newhamp$votesys=="D",]
5. DeanPerHand = newhamp[newhamp$Dean & newhamp$votesys=="H",]
7. DeanPerMach = newhamp[newhamp$Dean & newhamp$votesys=="D",]

10. plot(ObamaPerHand$pObama, DeanPerHand$Dean, pch = 6, xlab = "Obama Voters", ylab = "Dean Voters",
12. main = "Dean vs Obama Proportions", col="blue")
14. points(ObamaPerMach$pObama, DeanPerMach$Dean, pch = 1, col="orange")
16. legend("topleft", legend = c("Hand-Counted Votes", "Machine-Counted Votes"), pch = c(6, 1), col = c("blue", "orange"))

**Question 2**

1. **Normal Distribution**
2. (Code shown below)
3. The distribution shows the various degrees of freedom and how they differ.
4. x <- seq(-4, 4, length=1000)
5. y <- dnorm(x, mean=0, sd=1)
6. plot(x, y, type="l", lwd=1)
8. DF1 <- dt(x, 20)
9. lines(x, DF1, type="l", lwd=1, col="Red")
11. DF2 <- dt(x, 3)
12. lines(x, DF2, type="l", lwd=1, col="Green")
14. DF3 <- dt(x, 1)
15. lines(x, DF3, type="l", lwd=1, col="Blue")

**Question 3**

1. install.packages("Zelig")
2. library("Zelig")
3. data("voteincome")
4. ?voteincome
6. table(voteincome$year)
7. table(voteincome$income)
8. table(voteincome$age)/1500
10. sd(voteincome$age)
11. mean(voteincome$age)
12. **Null Hypothesis:** Average voter age is different than 50

**Alternative Hypothesis:** Average voter age is 50

1. Standard Error **= 0.47**; Z – Score = **-0.74;** P – Value = **0.23**
2. Failed to reject Null Hypothesis
3. Confidence Interval**: 48.73, 50.57**
4. D proves that C is indeed correct with a 95% confidence interval.
5. Confidence Interval Code:
6. z95 <- qnorm((1 - .95)/2, lower.tail = FALSE)
7. n <- length(na.omit(voteincome$age))
8. sample\_mean <- mean(voteincome$age, na.rm = TRUE)
9. sample\_sd <- sd(voteincome$age, na.rm = TRUE)
10. lower\_95 <- sample\_mean - (z95 \* (sample\_sd/sqrt(n)))
11. upper\_95 <- sample\_mean + (z95 \* (sample\_sd/sqrt(n)))
12. confint95 <- c(lower\_95, upper\_95)

**Question 4**

1. T – Test
2. For the example found in question 4, the T – Test would be used instead of the Z – Test for several reasons. One is that for small samples of under 30 observations, the T-distribution is used instead of the normal distribution. Another assumption would be whether the test is one-tailed or two-tailed. Another would be that the provided standard deviation is equal to the population standard deviation.
3. Test Statistic = **1.666667**; P Value **= .058156**; At alpha level 0.05, the result is **not significant** and the null can be rejected.
4. Assuming we know that the population standard deviation is equal to 1.2, we may be able to use a Z – test instead of a T – Test.
5. P Value is now **.048**
6. Null Hypothesis is not rejected
7. The difference between c) and f) is that the former requires a T – Test and the latter requires a Z – Test.

**Question 5**

1. If the voters are only given two options (Trump and Clinton) and the sample size is relatively large, the distribution should be relatively normal.
2. 341 (voters who prefer Trump)
3. Standard Error: **0.6745**
4. 95% Confidence Interval: **112, 115**

**Question 6**

1. In their conclusion, Green, Gerber, and Nickerson state that “mobilization campaigns have the potential to increase turnout substantially in local elections.” The authors found in their experiment that contacting registered voters via “door-to-door campaigns” significantly increases the likelihood that they will vote in elections.
2. The authors examine variables that are “manipulated” by an experimental study. Therefore, the treatment variables are whether they were contacted by the canvassers.
3. The outcome variable is the 7% increase that the study saw after successful canvassing.
4. The authors can claim that their findings are causal, because the study had a large effect on participants. Every time the potential voter was successfully contacted by the canvasser, that potential voter’s likelihood of voting in the election was raised by 7%, according to the authors.

**Question 7**

1. mean <- 1117
2. n <- 31
3. sample(mean, sd, replace = FALSE, prob = NULL)
4. z = (xbar - mean)/ (s/sqrt(n))
5. 2 \* pnorm(-abs(z))
6. pnorm(-abs(z))
7. Standard Error = **0.105**; T Score = **1.24**; P value **= 0.1075**; When alpha = .05, the given P value being greater than that shows the population means do not differ.
8. **No**, there would not be a 0 in the confidence interval, since the formula relates to only two averages.
9. **Normal Distribution**: , I would think that the distribution is approximately normal, since the number of observations it contains is quite large (1,117 for women and 870 for men).
10. mean = 2.86
11. s = 2.22
12. n = 870
13. xbar = 4.04
15. z = (xbar - mean)/ (s/sqrt(n))
16. 2 \* pnorm(-abs(z))
17. pnorm(-abs(z))

Alpha: 1.9678

To compare the two means, we subtract the m1 from m2.

**Question 8**

Standard Error = **0.898**; T - Score **= 0.145**; P - value = **2.060**;

P – value is incredibly high for the relatively small distribution, meaning that it is not useful.

**COLLABORATED WITH SINCLAIR BOWMAN AND LUKE EHERENSTROM**