STA371G Homework Assignment 6

(60 Points. Group homework.) Please write down the NAME and EID of each group member. Each group consists of up to three members. Each member needs to submit his/her own report

Problem 1 (5 points)

Read the report of Anzalone Liszt Grove Research on the NSA Surveillance Programs: "The public strongly supports... This memo is based on the results of a national poll conducted by Anzalone Liszt Grove Research. The poll of N=803 adults was conducted November 11-17, 2013. At least 30% of all interviews were conducted via cell phone and interviews were conducted in English and Spanish. The margin of error for the poll is plus or minus 3.5 percentage points at the 95% level of confidence."

(a) Can you explain why the margin of error for the poll is $\pm 3.5\%$ at the 95% level of confidence? (hint: the value of $p(1-p)$ is largest when $p = 0.5$)

(b) If Anzalone Liszt Grove Research conducts another poll of $N = 8,000$ adults, what would be the margin of error at the 95% level of confidence?

Problem 2 (5 points)

According to the official Federal Election Commission report for the presidential election in 2012, out of a total of 7,993,851 votes in Texas, President Barack Obama received 3,308,124 votes.

(a) If you randomly survey 1000 Texas residents who had voted in the 2012 presidential election, can you predict the distribution of the number of votes for President Obama among these 1000 Texas voters? Will you be surprised to find out that more than 500 of them voted for President Obama? (Hint: using the normal approximation to the binomial distribution.)

(b) According to http://www.politico.com/2012-election/results/president/texas/, In Dallas County, TX, President Obama received 57.1% of the votes in the 2012 presidential election. If you randomly survey 100 residents of Dallas County who had voted in the 2012 presidential election, can you predict the distribution of the number of votes for President Obama among them? Will you be surprised to find out no more than 50 votes for President Obama?

Problem 3 (50 points)

(a) Use R/Excel to simulate 1000 normal random numbers with mean 0.5 and standard deviation 0.6. Record the sample mean and sample variance of these 1000 simulated random numbers.

(b) Simulate a random number u that is uniformly distributed between 0 and 1, record its value and write it down:

(c) Calculate the probability that $X \sim \mathcal{N}(0.5, 0.6^2)$ is larger than u obtained in (b).

(d) Use these 1000 normal random numbers obtained in (a) to find an approximate answer to (c), record its value and write it down:

(e) Designing a simulation procedure to more accurately approximate $P(X > u)$.

(f) Find out that X is smaller than which number with probability u (note u is between 0 and 1 and hence can be treated as a probability).

(g) Using these 1000 normal random numbers obtained in (a) to find an approximate answer to (f), record its value and write it down:

(h) Given the 1000 random numbers obtained in (a), provide your 95% confidence interval of the true mean of the probability distribution that these 1000 random numbers are simulated from.

(i) To reduce the width of the 95% confidence interval of the true mean in (h) by about 50%, how many more random samples from the underlying distribution are needed?

(j) Simulate these additional random numbers and combing them with the original 1000 random numbers to find the sample mean and sample variance, and provide your updated 95% confidence interval of the true mean.