Discussion 6 Soln

1. A large mail-order club that offers monthly specials wishes to try out a new item. A trial mailing is sent to a random sample of 250 members selected from the list of over 9000 subscribers. Based on this sample mailing, 70 of the members decide to purchase the item. Give a 90% confidence interval for the proportion of club members that could be expected to purchase the item.

   Answer: $\hat{p} = \frac{70}{250} = 0.28$ *and* $SE(\hat{p} = \sqrt{\frac{.28*.72}{250}} = 0.028$, *so* $.28 \pm 1.645 * 0.028$=*(0.23394, 0.32606)*

2. A public health survey is to be designed to estimate the proportion p of a population having defective vision. How many persons should be examined if the public health commissioner wishes to be 98% certain that the margin of error is less than 0.05 when:

   (a) There is no knowledge about the value of p?

   Answer: Since p is unknown, use .5 for p, $ME = 0.05 = \sqrt{\frac{.5*.5}{n}}$ *so* $n = .5 * .5(\frac{2.33}{0.05})^2 = 543$

   (b) p is known to be about 0.3?

   Answer: $n = (.3 * .7)(\frac{2.33}{0.05})^2 = 456$

3. Generate qqplots when sampling from a Normal distribution for 10-15 different samples of size 10. Include the qqplots that shows the sample(s) that are most and least similar to what we would expect from a normal population. Make sure to label which is which and explain what characteristics of the graphs you used to choose which was which.

4. Consider a population having a discrete uniform distribution that places a probability of 0.1 on each of the integers 0,1,...,9. This may be an appropriate model for the distribution of the last digit in telephone numbers. Take 10000 samples of size 2, calculate the mean for each, and make a histogram to approximate the sampling distributio of $\bar{X}$. Do you think size 2 samples are large enough for the CLT to apply? If not, experiment with some different sample sizes to determine how large of a sample size is needed before you are comfortable that the CLT applies.

   something even close to 8 looks pretty good.

```
#Please help them set this up. Talk through a bit of what it does. If you think there is a cleaner way
pop1<-c(rep(0:9, each=1))
hist(pop1, freq=FALSE, breaks=seq(-.5, 12.5, by=1))

sampling.dist<-function(sample.size,pop, numExp=10000){
  estimate=rep(0, times=numExp)
  for (i in 1:numExp){
    sample<-sample(pop,sample.size, replace=TRUE)
    estimate[i]<-mean(sample)
  }
  return(estimate)
}

xbar.n2=sampling.dist(sample.size=2, pop=pop1, numExp=10000)
hist(xbar.n2, freq=FALSE, main="Approx. Sampling Distribution of X-Bar with n=2")
```

```r
qqnorm(xbar.n2)
mean(xbar.n2); var(xbar.n2)
```