Stat 140 - Quiz 3

What's Your Name? ____________________________

This is a sample quiz. On the real quiz, I will pick a different example of inference for a population proportion and ask questions similar to those below.

Garden Seeds

A garden center wants to store leftover packets of vegetable seeds for sale the following spring, but the center is concerned that the seeds may not germinate at the same rate a year later.

The manager randomly selects 200 seeds from a variety of packets from different companies and different batches of seeds, and plants them. She carefully prepares similar soil for the seeds in a greenhouse, and randomizes where the seeds are planted within the greenhouse. She finds that 171 of the 200 test seeds grew. The seed packets all claimed that 95% of the seeds would grow.

In answering the questions below, use the following output from R:

```r
binom.test(x = 171, n = 200, p = 0.95, alternative = "less")
```

```
##
##  Exact binomial test
##
## data:  171 and 200
## number of successes = 171, number of trials = 200, p-value =
## 2.948e-07
## alternative hypothesis: true probability of success is less than 0.95
## 95 percent confidence interval:
##  0.0000000 0.8942866
## sample estimates:
## probability of success
##                  0.855
```

```r
binom.test(x = 171, n = 200, p = 0.95, alternative = "greater")
```

```
##
##  Exact binomial test
##
## data:  171 and 200
## number of successes = 171, number of trials = 200, p-value = 1
## alternative hypothesis: true probability of success is greater than 0.95
## 95 percent confidence interval:
##  0.8075994 1.0000000
## sample estimates:
## probability of success
##                  0.855
```

```r
binom.test(x = 171, n = 200, p = 0.95, alternative = "two.sided")
```

```
##
##  Exact binomial test
##
## data:  171 and 200
## number of successes = 171, number of trials = 200, p-value =
## 2.948e-07
## alternative hypothesis: true probability of success is not equal to 0.95
## 95 percent confidence interval:
##  0.7984385 0.9006914
## sample estimates:
## probability of success
##                  0.855
```

1) Describe the population parameter in a sentence. What symbol would you use for the population parameter?

The population parameter is the proportion of seeds that germinate after being stored for a year. We will denote this parameter by the letter p.

2) Describe the sample statistic in a sentence. What symbol would you use for the sample statistic? What is the value of the sample statistic in this data set?

The sample statistic is the number of seeds that grew in the sample of 200 test seeds. In this case, we recorded $X = 171$ seeds that grew.

3) What are the null and alternative hypotheses for a test of the claim that after a year of storage, the seeds germinate at a lower rate than the package label indicates?

$H_0 : p = 0.95$

$H_A : p < 0.95$

4) What would you use for the sampling distribution of the sample statistic, if the null hypothesis for the test were true? Check all relevant conditions for using this sampling distribution.

Sampling distribution, if H_0 is true: $X \sim \text{Binomial}(200, 0.95)$

We need to check the following conditions:

Bias: There are no forms of bias in the sampling procedure. The seeds used were randomly selected.

Categorical variable: Yes, each seed either germinated or not.

Sample statistic is a count: Yes, we recorded 171 seeds that germinated.

Independent: The seeds were selected from different packets and different companies, so they . The seeds were all planted in the same greenhouse, so we might have concerns about independence; for example, if they left the windows too closed on a hot day, all of the seeds in the experiment could have been affected. Assuming they were careful in monitoring conditions in the greenhouse (which seems likely given that this is their business), I don't think this is a big deal.

Regardless of your answer about checking the conditions, let's go ahead with inference, for the sake of this quiz.

5) Draw a conclusion for the hypothesis test at the $\alpha = 0.05$ significance level. Explain what your answer means about germination rates of seeds after they've been stored for a year.

(The following sentence is my explanaation; you don't need to write this. Our alternative hypothesis is that $p < 0.95$, so we will use the output from the first call to `binom.test`.)

The p-value for the test is $2.948e - 07 = 2.948 * 10^{-7} = 0.0000002948$. This is less than the significance level of $\alpha = 0.05$, so we reject the null hypothesis. The data provide very strong evidence that seeds germinate at a lower-than-advertised rate after being stored for a year.

6) State a 95% confidence interval in the context of this problem. As part of your answer, explain what the phrase "95% confident" means.

(The following sentences are my explanation; you don't need to write this. We always use 2-sided confidence intervals. So we will use the confidence interval from the output from `binom.test` where the alternative was "two.sided".)

We are 95% confident that after being stored for a year, the proportion of seeds that will germinate is between about 0.798 and 0.901. If we were to take many different samples of seeds and use the results from each sample to compute a 95% confidence interval, about 95% of those confidence intervals would contain the true proportion of seeds that will germinate after being stored for a year.