## 10.4 Hypothesis Testing: Two Independent Samples Proportion

Example 3:

Smoking cigarettes has been known to cause cancer and other ailments. One politician believes that a higher tax should be imposed on those who smoke. An observational study was conducted in a suburb of Baltimore where 900 random people were asked the following question "Should the federal tax on cigarettes be raised to pay for health care reform?". Of those that were asked this question, 650 respondents were non-smokers (NS) and 250 respondents were smokers (S). 480 non-smokers said yes and 110 smokers said yes. Is there a statistical difference between the opinions of those who smoke and those who do not, with a significance level of 0.01?

1. Identify and State the Statistical Question
   - Determine the variable(s) of interest
   - Determine the type variable(s) (i.e., quantitative or qualitative)
   - Identify and state the hypotheses (Null and Alternative Hypotheses) based on the question at hand
2. Identify and state level of significance α (the probability of rejecting the H_0 when H_0 is true)
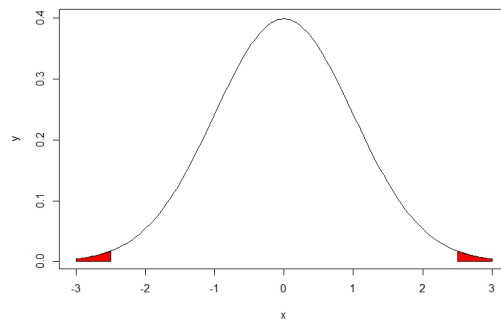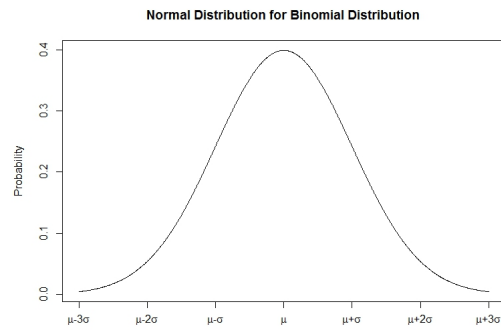

Really IMPORTANT:
   - Critical Value:
   - α:

3. Perform Statistical Test and Interpret Results

$$TS = z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}}$$

Normal Distribution for Binomial Distribution

- Test Statistic:
- p-value:

4. State the sample, null hypothesis, test that was used, and conclusion with non-statistical terms
   -
   -
   -
   -

10.5 Hypothesis Testing: Two Independent Samples Standard Deviation

We did not do a hypothesis test for one sample standard deviation due to time, however, understanding the process for hypothesis testing of two groups standard deviations (variation) is critical to a concept discussed later this quarter.

Review Assumptions:

1. Independent Samples
2. Two populations are normally distributed

Note: The standard deviation follows a Chi-Squared Distribution. Some small facts:

- The standard deviation is always positive $\rightarrow$ Chi-square Distribution is always positive
- This distribution requires degrees of freedom which is $n - 1$

What other distribution requires degrees of freedom?

Example 4:
Researchers conducted a study to determine whether Treatment 1 works better than Treatment 2 in pain reduction. Use a significance level 0.05 to test the the claim that those that had Treatment 1 vary more than those that had Treatment 2. **Process for Hypothesis**

Table 3: Summary of Data

	n	$\bar{x}$	s
Treatment 1	20	0.44	1.4
Treatment 2	20	0.49	0.96

Testing for this class:

1. Identify and State the Statistical Question
   - Determine the variable(s) of interest
   - Determine the type variable(s) (i.e., quantitative or qualitative)
   - Identify and state the hypotheses (Null and Alternative Hypotheses) based on the question at hand

2. Identify and state level of significance α (the probability of rejecting the H_0 when H_0 is true)
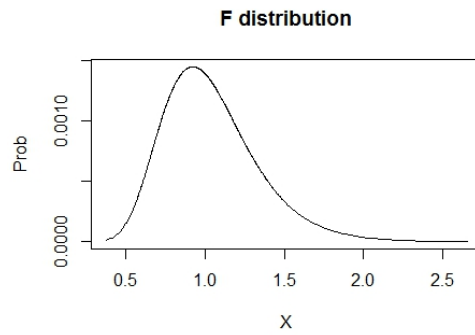
F distribution
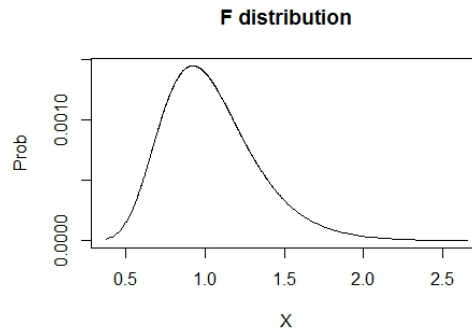
Really IMPORTANT:
   - Critical Value, using $df_1 = n_1 - 1, df_2 = n_2 - 1$:
   - α:

3. Perform Statistical Test and Interpret Results

$$TS = F = \frac{s_1^2}{s_2^2}$$

F distribution

- Test Statistic:
- p-value:

4. State the sample, null hypothesis, test that was used, and conclusion with non-statistical terms
  -
  -
  -
  -

10.6   Confidence Intervals Revisited

Review of a Confidence Interval (CI)

- PE: Point Estimate
- ME: Margin of Error
- CV: Critical Value, based $1 - \alpha$ confidence level
- SE: Standard Error
- $PE \pm ME = PE \pm CV(SE)$
- $LB = PE - ME = PE - CV(SE)$
- $UB = PE + ME = PE + CV(SE)$
- We are $1 - \alpha$ % confident that the true parameter is between LB and UB
- **NOTE:** Remember a CI is an estimate and gives the same result as an two-tailed hypothesis test.
- If 0 is within the LB and UB and you are testing if a parameter equal 0, you will always fail to reject the null hypothesis. *Remember this for next week.*

10.6.1   CI: Two Independent Samples Mean

Example 5:
Monterey and Santa Cruz Trees.

	$\bar{x}$	s	n
Santa Cruz	310.36	32.63	27
Monterey	269.23	30.3	24

Create a 86 % CI.
Process:

1. $PE = (\bar{x}_1 - \bar{x}_2) =$

2. $SE = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} =$

3. $1 - \alpha\% = \qquad \alpha =$

4. $CV = t_{\alpha/2} = \qquad$ from t-table using $\alpha/2, df =$

5. $LB = PE - (CV)SE =$

6. $UB = PE + (CV)SE =$

7. **Interpretation:** We are $1 - \alpha\%$ confident that the difference between population means is within the $LB = \qquad$ and $UB = \qquad$.

10.6.2   CI: Two Dependent (Paired) Samples Mean

Example 6:
GPA Improvement Program
$\bar{x}_d = 0.3014, \quad s_d = 0.5065$ Create a 94 % CI.
Process:

1. $PE = (\bar{x}_d) =$

2. $SE = s_d/\sqrt{n} =$

3. $1 - \alpha\% = \quad \alpha =$

4. $CV = t_{\alpha/2} = \quad$ from t-table using $\alpha/2, df =$

5. $LB = PE - (CV)SE =$

6. $UB = PE + (CV)SE =$

7. **Interpretation:** We are $1 - \alpha\%$ confident that the mean difference between populations is within the $LB =$ and $UB =$.