1. The distribution of scores of students taking the LSATs is claimed to have a mean of 521. We take a sample of 100 incoming Harvard Law School freshman LSAT scores and find a mean of 589 and a standard deviation of 37. Since Harvard is an Ivy League school, they think their freshmen are smarter than average law students. Test this theory (that Harvard students score higher than average on the LSATs) at the 0.05 significance level.

Ho: \( \mu = 521 \)  
Ha: \( \mu > 521 \)

Conditions:
1. Random – Not stated so we will assume that the sample is representative of the all Harvard freshman
2. 10% Condition – 100 Harvard freshman is less than 10% of all Harvard freshman
3. Nearly Normal – no graph is given but sample size(100) is much larger than 30 so the CLT applies

All conditions have been met to use Student’s t-model for a 1 sample t-test.

\[ t_{99} = \frac{\bar{x} - \mu}{s / \sqrt{n}} = \frac{589 - 521}{37 / \sqrt{100}} = 18.378 \]

P-Value = \( P(t_{99} > 18.378) = 5.6007 \times 10^{-34} \)

Since the P-Value is less than alpha (5.6007 \( \times 10^{-34} < 0.05 \)) we reject the null hypothesis. There is statistically significant evidence that the mean LSATS for Harvard freshman is higher than the average LSATS.
2. A teacher wants to test the effectiveness of a new textbook. She believes that this new textbook is easier to read, and that her students should have better grades on their tests this year than they have in the past. She took a random sample of test scores from last year’s classes, and then a random sample of test scores from this year’s classes. Assume normal populations for both years. Test her theory at $\alpha = 0.01$. 

<table>
<thead>
<tr>
<th>Old book</th>
<th>New book</th>
</tr>
</thead>
<tbody>
<tr>
<td>85 84 91 75 65</td>
<td>94 62 86 89 80</td>
</tr>
<tr>
<td>75 82 84 89 62</td>
<td>96 88 88 79 75</td>
</tr>
<tr>
<td>74 64 58 95 50</td>
<td>94 84 86 78 64</td>
</tr>
</tbody>
</table>

$\mu_O = \text{mean test score from class using the old book}$

$\mu_N = \text{mean test score from class using the new book}$

Hypothesis:

H$_0$: $\mu_O - \mu_N = 0$; There is no difference in the mean test scores from last year using the old book and this year using the new book.

H$_A$: $\mu_O - \mu_N < 0$; The mean test scores from last year are lower than the mean scores from this year.

Conditions:

1) Randomization: Stated as random samples
2) 10% Condition: 15 students is less than 10% of all students that have used either textbook.
3) Independence: The two classes would be independent of each other
4) Nearly Normal: Histograms are unimodal and approximately symmetric

All conditions have been met to use Student’s $t$-model for a two sample $t$-test.

Mechanics:

$n_O = 15$, $\bar{x}_O = 75.53$, $s_O = 13.27$, $\alpha = 0.01$

$n_N = 15$, $\bar{x}_N = 82.87$, $s_N = 10.11$

$df = 26.2$

$t_{26} = \frac{75.53 - 82.87}{\sqrt{\frac{13.27^2}{15} + \frac{10.11^2}{15}}} = -1.703$

$P$-Value = $P(t_{26} < -1.703) = 0.0502$

Conclusion:

Since the $P$-Value is greater than alpha ($0.0502 > 0.01$) we fail to reject the null hypothesis. There is not enough evidence to state that the mean scores from last year were less than this year. There is not enough evidence to state that the students perform better with the new textbook.
3. A football coach is frustrated with his team’s lack of speed. He measures each player’s 40-yard dash speed and then sends all of them to a speed and agility camp. He then measures their times again after. The data is below. Is their sufficient evidence to say that the camp helped the players speed? Run a test.

<table>
<thead>
<tr>
<th>Before</th>
<th>4.88</th>
<th>5.10</th>
<th>4.41</th>
<th>4.73</th>
<th>4.60</th>
<th>4.86</th>
<th>4.95</th>
<th>4.98</th>
<th>5.10</th>
<th>5.13</th>
<th>5.05</th>
<th>4.90</th>
<th>4.70</th>
<th>4.60</th>
<th>5.11</th>
</tr>
</thead>
<tbody>
<tr>
<td>After</td>
<td>4.70</td>
<td>4.85</td>
<td>4.35</td>
<td>4.77</td>
<td>4.56</td>
<td>4.78</td>
<td>4.70</td>
<td>4.90</td>
<td>5.10</td>
<td>5.10</td>
<td>4.70</td>
<td>4.56</td>
<td>4.34</td>
<td>4.90</td>
<td></td>
</tr>
</tbody>
</table>

$\mu_d = \mu_{\text{Before}} - \mu_{\text{After}}$

Hypothesis:
$H_0: \mu_d = 0$; The mean difference in time before the camp and after the camp is zero.
$H_A: \mu_d > 0$; The mean difference in time before the camp and after the camp is greater than zero. There has been a reduction in the time of a player’s 40-yard dash.

Conditions:
1) Paired Data: Times were recorded on the same players before and after the camp.
2) Randomization: Not stated as a random sample but we will assume that the sample is representative of all football players
3) 10% Condition: 15 football players is less than 10% of all football players.
4) Nearly Normal: The normal probability plot show an approximately linear pattern

All conditions have been met to use Student’s $t$-model for a matched pairs $t$-test.

Mechanics:
$\bar{x}_d = 0.122$  $s_d = 0.108$  $\alpha = 0.05$
$n = 15$  $df = 14$
$t_{14} = \frac{0.122 - 0}{0.108} = 4.387$

P-Value = $P(t_{14} > 4.387) = 3.103 \times 10^{-4}$

Conclusion:
Since the P-Value is less than alpha ($3.103 \times 10^{-4} < 0.05$) we reject the null hypothesis. There is statistically significant evidence that the mean difference in times from before the camp to after the camp has decreased. It appears that the camp has helped the player’s speed.
4. Poisoning by DDT causes tremors and convulsions and slows recovery times of muscles. In a study of DDT poisoning, researchers fed several lab rats a measured amount of DDT. They then made measurements of the rats’ refractory period (the time needed for a nerve to recover after a stimulus). In their sample they find the following times: 1.61, 1.9, 1.53, 1.4, 1.33, 1.81, 1.3, 1.25, 1.65.

a. Estimate the average refractory period using 95% confidence.

Conditions:
1) Randomization: Not stated as a random sample but we will assume that the sample is representative of all rats
2) 10% Condition: 9 rats is less than 10% of all rats
3) Nearly Normal: The normal probability plot shows an approximate linear pattern

All conditions have been met to use Student’s $t$-model for a one sample $t$-interval.

$$\bar{x} = 1.531 \quad s = 0.23 \quad n = 9$$

$$df = 8 \quad t_8^* = 2.306$$

$$1.531 \pm 2.306 \frac{0.23}{\sqrt{9}} = (1.3541, 1.7081)$$

We are 95% confident that the true average refractory period is between 1.3541 and 1.7081 milliseconds.

b. If we know that the mean time for unpoisoned rats is 1.3 milliseconds, does your interval give evidence that the average time is different for poisoned rats?

Yes, our interval gives us evidence that the mean time for poisoned rats is higher than unpoisoned because the entire 95% confidence interval is above 1.3 milliseconds. We are 95% confident that the average refractory time for poisoned rats is above 1.3 milliseconds.
5. The Chapin Social Insight Test is a psychological test designed to measure how accurately a person appraises other people. The possible scores on the test range from 0 to 41. During the development of the test, it was given to several groups of people. Here are the results for male and female college students at a liberal arts college:

<table>
<thead>
<tr>
<th></th>
<th>n</th>
<th>avg.</th>
<th>std.dev</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>133</td>
<td>25.34</td>
<td>5.05</td>
</tr>
<tr>
<td>Female</td>
<td>162</td>
<td>24.94</td>
<td>5.10</td>
</tr>
</tbody>
</table>

Does the data support the contention that female and male students differ in average social insight? Use 96% confidence to make your conclusion.

\[ \mu_1 = \text{mean social insight for males} \]
\[ \mu_2 = \text{mean social insight for females} \]

Conditions:
1) Randomization: Not stated as random so we will assume that each sample is representative of the two populations
2) 10% Condition: 133 males is less than 10% of all male college students. 162 females is less than 10% of all female college students
3) Independence: Males and females would be independent of each other
4) Nearly Normal: No graphs were given but the sample sizes are both large (n > 30) so the CLT applies

All conditions have been met to use Student’s \( t \)-model for a two sample \( t \)-interval.

\[ df = 283 \quad t_{283} = 2.063 \]

\[ (25.34 - 24.94) \pm 2.063 \sqrt{\frac{5.05^2}{133} + \frac{5.10^2}{162}} = (-0.8247, 1.6247) \]

We are 96% confident that the average score of men is between 0.8247 lower to 1.6247 points higher than that of women. Since 0 is in the interval, we have evidence to say that there is no difference between the scores of the two genders.
6. Many drivers of cars that can run on regular gas actually buy premium in the belief that they will get better gas mileage. To test that belief, we use 10 cars in a company fleet in which all the cars run on regular gas. Each car is filled first with either regular or premium gasoline, decided by a coin toss, and the mileage for that tank-full is recorded. Then the mileage is recorded again for the same cars for a tank-full of the other kind of gasoline. We don’t let the drivers know about this experiment. Here are the results in miles per gallon:

<table>
<thead>
<tr>
<th>Regular</th>
<th>16</th>
<th>20</th>
<th>21</th>
<th>22</th>
<th>23</th>
<th>22</th>
<th>27</th>
<th>25</th>
<th>27</th>
<th>28</th>
</tr>
</thead>
<tbody>
<tr>
<td>Premium</td>
<td>19</td>
<td>22</td>
<td>24</td>
<td>24</td>
<td>25</td>
<td>25</td>
<td>26</td>
<td>26</td>
<td>28</td>
<td>32</td>
</tr>
</tbody>
</table>

Is there evidence that cars get significantly better fuel economy with premium gasoline? Use 0.01 level of significance and a test.

\[ \mu_d = \mu_{\text{regular-premium}} \]

Hypothesis:
- \( H_0: \mu_d = 0; \) The mean difference of mpg between regular gasoline and premium gasoline is zero.
- \( H_A: \mu_d < 0; \) The mean difference of mpg between regular gasoline and premium gasoline is less than zero.

Conditions:
1) Paired data: Data is from individual cars using both regular and premium gasoline.
2) Randomization: Each car is randomly assigned to a treatment
3) Independence: Each car’s mpg would be independent of each other
4) Nearly Normal: The normal probability plot of the differences is approximately linear with no large outliers.

All conditions have been met to use Student’s \( t \)-model for a matched pairs \( t \)-test.

Mechanics:
\[
\bar{x}_d = -2 \quad s_d = 1.414 \quad \alpha = 0.05 \\
\begin{align*}
n &= 10 \\
df &= 9 \\
t_9 &= \frac{-2 - 0}{\frac{1.414}{\sqrt{10}}} = -4.472 \\
P\text{-Value} &= P(t_9 < -4.472) = 7.749 \times 10^{-4}
\end{align*}
\]

Conclusion:
Since the P-Value is less than alpha (7.749 \times 10^{-4} < 0.05) I would reject the null hypothesis. There is statistically significant evidence that mean difference of mpg between regular and premium gasoline is less than zero. It appears that cars get better mpg using premium gasoline.
7. For inference on means, why is Student’s t-model used and not the Normal model? Because we do not know both the mean and the standard deviation of the population. Therefore we cannot use Z-scores and cannot use the normal model. Student’s t-model is used when the population standard deviation is not known and must be estimated using the sample standard deviation. Estimating the standard deviation creates more variability.

8. How is Student’s t-model the same as the Normal model? They are both: unimodal, symmetric, centered at 0.

9. How is Student’s t-model different than the Normal model? The student’s t-model is wider than the normal model.

10. How is degree of freedom calculated and what effect does it have on the Student’s t-model? 
   df = n – 1 for 1 sample procedures 
   df = on the calculator for 2 sample procedures 
   As the df increases, the t-distribution becomes narrower, and closer to the normal model.

11. A certain population is skewed to the right. We want to estimate the mean so we take a sample. What must we know about the sample if we wish to create a confidence interval to estimate the true mean? We must know that the sample size is greater than or equal to 30, otherwise our third condition would not check out, and we could not do the confidence interval.

12. If we calculate a 95% confidence interval, how can we decrease the margin of error without losing confidence? Increase sample size.

13. What is the critical value (t*) for a sample of 67 for 92% confidence? 
   \( t_{66}^* = 1.778 \) (use INVT program)

14. What is the critical value for a sample of size 153 for 97% confidence? 
   \( t_{152}^* = 2.191 \)

15. What is the p-value if I have a test statistic of \( t = 2.145 \), sample size of 28, and am doing a 2-sided test? 
   P-value = \( 2P(t_{27} > |2.145|) = 0.0411 \)

16. What is the p-value if I have a test statistic of \( t = -1.987 \), sample size of 193, and am doing a lower-tailed test? 
   P-value = \( P(t_{192} < -1.987) = 0.0242 \)

17. I have an interval that is (102, 105).
   a. What is the sample mean? 103.5 units
   b. What is the margin of error? 1.5 units
   c. Assuming that the standard deviation is 5 and the sample size is 40, what is the confidence level? 
   \[ 1.5 = t_{99}^* \left( \frac{5}{\sqrt{40}} \right) \]
   \( t_{99}^* = 1.897 \)
   Conf level = tcdf(-1.897, 1.897, 39) = 93.47% confidence
We wish to see if the dial indicating the oven temperature for a certain model oven is properly calibrated. 12 ovens of this model are selected at random. The dial on each is set to 300° F; after one hour, the actual temperature of each is measured with a thermometer. The temperatures measured had a mean of 302.5° F with a standard deviation of 0.25° F. Assuming that the actual temperatures for this model when the dial is set to 300° are normally distributed, we test whether the dial is properly calibrated by testing the hypotheses $H_0: \mu = 300, H_a: \mu \neq 300$

18. Assuming conditions are met, calculate the t-statistic and P-Value for this problem.
\[ t_{11} = 34.641 \quad P\text{-Value} = 2P(t_{11} > 34.641) = 1.39 \times 10^{-12} \]

19. What does the P-Value mean in this context?
There is a $1.39 \times 10^{-10}$% chance of getting a sample where the mean temperature is 302.5° F or more extreme, if the mean temperature of the oven really is 300° F.

20. What would be a Type I error in this context?
We conclude that the oven temperature is not 300° F, when really it is 300° F.

21. What would be a Type II error in this context?
We conclude that the oven temperature is 300 ° F, when really it is not 300° F.

22. What is Power in this context?
We conclude that the temperature of the oven is not 300 ° F, and we are correct that it is not 300° F.

23. If a 96% confidence interval were calculated what would be the critical value?
\[ N = 12 \quad df = 11 \quad \text{Conf} = 96\% \]
\[ t^*_{11} = 2.328 \]

24. What does 96% confidence mean in this context?
About 96% of random samples of size 12 will produce confidence intervals that contain the true mean temperature of the ovens.

25. If the test was redone with the same significance level on 24 ovens and the same mean and standard deviation were found:
   a. What would happen to the P-Value?
      Decrease (because your test stat would increase)
   b. What would happen to the Type I error?
      Stay the same (because your alpha, or significance level would stay the same)
   c. What would happen to the Type II error?
      Decrease (opposite of power)
   d. What would happen to the Power?
      Increase (because n increases)
   e. How would the 96% confidence interval change?
      Decrease/smaller/skinnier (when the sample size increases, the margin of error decreases)

 NOTE: Extra Credit will be offered. Problems will be from Unit 5 material (proportions)