1. The weights of 39 boxes of cereal are measured. The sample standard deviation is 0.37 ounces. The sample mean is 15.9 ounces. We are concerned that the mean weight of all boxes may be under 16 ounces.

(a) State the null and alternative hypotheses. \( H_0 : \mu = 16, \quad H_1 : \mu < 16 \)

(b) Test the null hypothesis at the 5% level, using either p-value or critical value approach, showing calculation of the test statistic.

\[
t = \frac{\bar{x} - \mu}{s / \sqrt{n}} = \frac{15.9 - 16}{(0.37 / \sqrt{39})} = \frac{-0.1 \times \sqrt{39}}{0.37} = -1.68783729686
\]

p-value is \( \text{tcdf}(-10^{99}, -1.68783729686, 38) = 0.0498174749 \)

p-value < 0.05 so reject \( H_0 \). There is enough evidence at the 5% level to say the mean weight of all the boxes is under 16 ounces.

The critical value is \( t(0.95, 38) = \text{invT}(0.05, 38) = -1.6859544047 \)

The computed t-value above is just under this left hand critical value, which makes the computed value more extreme, so we reject the null hypothesis.

Can solve \( 0 = 0.050 - \text{tcdf}(-10^{99}, X, 38) \) to get the critical value.

2. The weights of 21 loaves of bread are measured. The sample standard deviation is 0.6 ounces. We are concerned that the weights of the loaves may have a standard deviation greater than 0.5 ounces.

(a) State the null and alternative hypotheses. \( H_0 : \sigma = 0.5, \quad H_1 : \sigma > 0.5 \)

(b) Test the null hypothesis at the 10% level, using either p-value or critical value approach, showing calculation of the test Statistic.

\[
\chi^2 = \frac{(n - 1)s^2}{\sigma^2} = \frac{20(0.6)^2}{(0.5)^2} = 28.8
\]

p-value is \( \chi^2 \text{cdf}(28.8, 10^{99}, 20) = 0.0917727417 < 0.10 \) so reject \( H_0 \). There is enough evidence at the 10% level to say the S.D. of the loaves is over 0.5 ounces.

Solve \( 0 = 0.1 - \chi^2 \text{cdf}(X, 10^{99}, 20) \) to get the critical value \( \chi^2(0.10, 20) = 28.411980585 \)

The computed value is more extreme, so we reject the null hypothesis.

GO TO PAGE 2.
3. In the previous election, 69% of the registered voters in Midville actually voted in the election for mayor. This year, a survey of 96 randomly chosen registered voters, found that 62 had voted in the current election for mayor. Has there been a significant decrease in the proportion of registered voters actually voting, at the 5% level? Show calculation of the test Statistic. 

\[ H_0 : p = 0.69, \quad H_1 : p < 0.69 \]

\[ \hat{p} = \frac{62}{96} = 0.645833 \]

\[ z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}} = \frac{0.62 - 0.69}{\sqrt{0.69(1-0.69)/96}} = -0.9356746164 \]

p-value is \( \text{normalcdf}(-10^99, -0.9356745164) = 0.1747203519 > 0.05 \). Fail to reject \( H_0 \).

There is not enough evidence at the 5% level to say that the proportion voting has decreased.

Assuming a large population we can use binomial probability. The p-value is \( \text{binomcdf}(96, 0.69, 62) = 0.2033477978 > 0.05 \). Again, fail to reject \( H_0 \).

4.

<table>
<thead>
<tr>
<th></th>
<th>Student</th>
<th>Carla</th>
<th>Rose</th>
<th>Kris</th>
<th>Dean</th>
<th>Alberto</th>
</tr>
</thead>
<tbody>
<tr>
<td>( x_1 )</td>
<td>Test One</td>
<td>78</td>
<td>70</td>
<td>82</td>
<td>92</td>
<td>81</td>
</tr>
<tr>
<td>( x_2 )</td>
<td>Test Two</td>
<td>79</td>
<td>75</td>
<td>81</td>
<td>90</td>
<td>86</td>
</tr>
<tr>
<td>( d = x_2 - x_1 )</td>
<td>1</td>
<td>5</td>
<td>-1</td>
<td>-2</td>
<td>5</td>
<td></td>
</tr>
</tbody>
</table>

Test the hypothesis that there is no significant improvement in average score from test one to test two, based on the above sample of five students, at the 1% level of significance. Show calculation of the test statistic. 

\[ H_0 : \mu_d = 0, \quad H_1 : \mu_d > 0 \]

Use 1-Var Stats function on the TI to get mean and standard deviation for the differences.

\[ \bar{d} = 1.6, \quad s_d = 3.286335345, \quad t = \frac{\bar{d} - 0}{s_d / \sqrt{n}} = \frac{1.6}{3.286335345 / \sqrt{5}} = 1.088662108 \]

p-value is \( \text{tcdf}(1.088662108, 10^99, 4) = 0.1687509283 > 0.01 \) Fail to reject \( H_0 \).

There is not enough evidence at the 1% level to say there is a significant improvement.

Note that we could have used \( d = x_1 - x_2 \) and \( H_0 : \mu_d = 0, \quad H_1 : \mu_d < 0 \).
5. One class, of 35 students, averages 77 on a test with variance, \( s_1^2 = 16 \). Another class, of 31 students, averages 81 on the same test, with variance, \( s_2^2 = 25 \).

(a) Compute the estimated standard error of the difference of sample means.  
(You have to identify exactly what this is, here.)

\[
\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = \sqrt{\frac{16}{35} + \frac{25}{31}} = \sqrt{1.26359447} = 1.12409718 \text{ (Store this value.)}
\]

(b) Are the averages significantly different at the 2% level? Show calculation of the test statistic. May use TI to get the degrees of freedom. State hypotheses.

\[
H_0 : \mu_1 = \mu_2, \quad H_1 : \mu_1 \neq \mu_2 \quad df = 57.38205774
\]

\[
t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{77 - 81}{1.12409717998...} = -3.558411204
\]

p-value is \( 2\text{tcdf}(-10^{99}, -3.558411204, 57.38205774) = 0.0007568395 < 0.02 \)

Reject \( H_0 \). There is enough evidence at the 2% level to say the population means differ. (Assuming these classes are simple random samples from normal populations.)

6. As in the previous problem, one class, of 35 students, averages 77 on a test with variance, \( s_1^2 = 16 \). Another class, of 31 students, averages 81 on the same test, with variance, \( s_2^2 = 25 \). Is the sample variance of the 2nd class significantly higher than that of the first class, at the 10% level? Show calculation of the test statistic. (Assume these classes are simple random samples from normal populations.) State hypotheses.

\[
H_0 : \sigma_1 = \sigma_2, \quad H_1 : \sigma_1 < \sigma_2 \quad F = \frac{s_1^2}{s_2^2} = \frac{16}{25} = 0.64
\]

p-value is \( \text{Fcdf}(0, 16/25, 34, 30) = 0.1041819866 \)

or, \( F = \frac{s_2^2}{s_1^2} = \frac{25}{16} = 1.5625 \) p-value is \( \text{Fcdf}(25/16, 10^{99}, 30, 34) = 0.1041819866 \)

p-value > 0.10 so fail to reject \( H_0 \).

There is not enough evidence at the 10% level to say the variance (or standard deviation) of the 2nd class is significantly higher than that of the first class.
7. In a test of drugs A, B and C, the following results were obtained.

<table>
<thead>
<tr>
<th></th>
<th>Worse</th>
<th>Same</th>
<th>Improved</th>
</tr>
</thead>
<tbody>
<tr>
<td>Drug A</td>
<td>12</td>
<td>15</td>
<td>28</td>
</tr>
<tr>
<td>Drug B</td>
<td>15</td>
<td>18</td>
<td>22</td>
</tr>
<tr>
<td>Drug C</td>
<td>12</td>
<td>20</td>
<td>25</td>
</tr>
<tr>
<td></td>
<td>39</td>
<td>53</td>
<td>75</td>
</tr>
</tbody>
</table>

(a) Give computed value of the appropriate statistic, and the p-value.
\[
\begin{align*}
\text{df} &= (#\text{rows} - 1)(#\text{cols} - 1) = (3 - 1)(3 - 1) = 2(2) = 4 \\
\chi^2 &= \sum \frac{(O - E)^2}{E} = 1.860505216 \\
p\text{-value} &= \chi^2\text{cdf}(1.860505216, 10^{99}, 4) = 0.7613959695
\end{align*}
\]

(b) Give the critical value of the test statistic, if \( \alpha = 0.05 \).
\[
\text{Solve } 0 = 0.05 - \chi^2\text{cdf}(X, 10^{99}, 4) \text{ to get } \chi^2(.05, 4) = 9.48772903678
\]

(c) Test the hypothesis that drug and outcome are independent, at the 5% level.
\[
\begin{align*}
\text{H}_0 & : \text{Drug,} & & \text{& result are independent, } \quad \text{H}_1 : \text{Drug,} & & \text{& result are dependent} \\
p &= 0.7613959695 > 0.05 = \alpha, \text{ so fail to reject } \text{H}_0.
\end{align*}
\]

There is not enough evidence at the 5% level to say drug and result are dependent.

(d) Find the expected frequency for Drug C - Improved, showing the work.
\[
\frac{75 \cdot 57}{167} = 25.5988024
\]

GO TO PAGE 5.
8. (a) Complete the table below and find

(b) the equation of the line of best fit in the form \( y = ax + b \) with formulas needed to find it

\[
\begin{align*}
SS(x) &= 110 - \frac{20^2}{4} = 10 \\
SS(y) &= 219 - \frac{29^2}{4} = 8.75 \\
SS(xy) &= 136 - \frac{20 \cdot 29}{4} = -9 \\
a &= \frac{SS(xy)}{SS(x)} = -\frac{9}{10} \\
\sum y &= a \sum x + nb \\
29 &= 20a + 4b = 20(-0.9) + 4b, \quad 4b = 47, \quad b = 11.75.
\end{align*}
\]

Can also use \( \sum xy = a \sum x^2 + b \sum x \), \( 136 = 110a + 20b \), with either of the above two equations, to get \( y = -0.9x + 11.75 \), the equation of the line of best fit.

(c) \( r \), the coefficient of linear correlation with formulas needed to find it

\[
r = \frac{SS(xy)}{\sqrt{SS(x) \cdot SS(y)}} = \frac{-9}{\sqrt{10(8.75)}} = -0.9621404709
\]

(d) the predicted \( y \) if \( x = 10 \). \( y = -0.9(10) + 11.75 = 2.75 \)

(e) Use a TI function to test \( H_0 : \rho = 0 \), vs \( H_1 : \rho \neq 0 \). Need only give \( t \) and \( p \) values.

Use LinRegTTest to get \( t = -4.992301766, \quad p = 0.0378595291 \)

\begin{center}
\begin{tabular}{|c|c|c|c|c|c|}
\hline
x & y & x^2 & xy & y^2 \\
\hline
3 & 9 & 9 & 27 & 81 \\
4 & 8 & 16 & 32 & 64 \\
6 & 7 & 36 & 42 & 49 \\
7 & 5 & 49 & 35 & 25 \\
20 & 29 & 110 & 136 & 219 \\
\hline
\end{tabular}
\end{center}
9. A researcher tests drugs A, B and C, and uses a quantitative measure of improvement. Three groups of people tried the drugs with results below.

(a) Use ANOVA and make an appropriate **table**

(b) Give the **computed** value of the appropriate statistic, and the **p-value**.

(c) Find the **critical value** of the test statistic, if $\alpha = 0.02$.

(d) Test the hypothesis that there is no difference in the mean improvement among the drugs, at the 2% level.

<table>
<thead>
<tr>
<th>Drug A</th>
<th>Drug B</th>
<th>Drug C</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>12</td>
<td>6</td>
</tr>
<tr>
<td>10</td>
<td>9</td>
<td>8</td>
</tr>
<tr>
<td>8</td>
<td>10</td>
<td>6</td>
</tr>
<tr>
<td>9</td>
<td>11</td>
<td>7</td>
</tr>
<tr>
<td>7</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**H₀**: Means are all equal.  **H₁**: Means are not all equal.

(a) Use ANOVA($L₁, L₂, L₃$)

<table>
<thead>
<tr>
<th></th>
<th>df</th>
<th>SS</th>
<th>MS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Factor</td>
<td>2</td>
<td>30.95</td>
<td>15.475</td>
</tr>
<tr>
<td>Error</td>
<td>12</td>
<td>18.783</td>
<td>1.56527</td>
</tr>
<tr>
<td>Total</td>
<td>14</td>
<td>49.73</td>
<td></td>
</tr>
</tbody>
</table>

(b) $F = \frac{MS(\text{factor})}{MS(\text{error})} = \frac{15.475}{1.56527} = 9.886424135$

The **p-value** is $F_{cdf}(9.886424135, 10^{99}, 2, 12) = 0.0029023553 \lessdot 0.02 = \alpha$

(c) Solve $0 = 0.02 - F_{cdf}(X, 10^{99}, 2, 12)$.  $F(0.02, 2, 12) = 5.516298622$

(d) $p < \alpha$.  **Reject** $H₀$.  There is enough evidence at the 2% level to say the means are not all equal.