Calculate Confidence Intervals Using the TI Graphing Calculator

Confidence Interval for Population Proportion \( p \)

Select: STAT / TESTS / 1-PropZInt
- \( x \): number of successes found in sample
- \( n \): sample size
- C-Level: 0.90, 0.95, 0.99, etc
Calculate: Select Calculate and press Enter

Program Output:
- Confidence Interval: (lower bound, upper bound)
- Calculated value of p-hat statistic
- Sample size \( n \)

Confidence Interval for Population \( \mu \) (\( \sigma \) is known)

Select: STAT / TESTS / ZInterval
- Inpt: Use arrow keys – select Stats
- \( \bar{x} \): value of sample mean statistic
- \( sx \): value of sample standard deviation statistic
- \( n \): sample size
- C-Level: 0.90, 0.95, 0.99, etc
Calculate: Select Calculate and press Enter

Program Output:
- Confidence Interval: (lower bound, upper bound)
- Value of x-bar statistic
- Sample size \( n \)

Confidence Interval Population \( \mu \) (\( \sigma \) is unknown)

Select: STAT / TESTS / TInterval
- Inpt: Use arrow keys – select Stats
- \( \bar{x} \): value of the sample mean statistic
- \( sx \): value of sample standard deviation statistic
- \( n \): sample size
- C-Level: 0.90, 0.95, 0.99, etc
Calculate: Select Calculate and press Enter

Output:
- Confidence Interval: (lower bound, upper bound)
- Value of the sample x-bar statistic
- Value of the standard deviation statistic \( s \)
- Sample size \( n \)

Confidence Interval Population \( \sigma \)
The current versions of the TI graphing calculators do not have a program to calculate confidence intervals for \( \sigma^2 \) and \( \sigma \). There is a program named S2INT that can be installed on a TI. See page 383 of your textbook for details.

CI for Difference of Population Means (\( \mu_1 - \mu_2 \))

Population \( \sigma \)'s are unknown, but we assume \( \sigma_1 = \sigma_2 \).
Sample statistics are taken from two independent simple random samples.

Select: STAT / TESTS / 2-SampTInt
- Inpt: Use arrow keys – select Stats
- \( \bar{x}_1 \): sample mean statistic of first sample
- \( sx_1 \): sample standard deviation of first sample
- \( n_1 \): sample size of first sample
- \( \bar{x}_2 \): sample mean statistic of second sample
- \( sx_2 \): sample standard deviation of second sample
- \( n_2 \): sample size of second sample
- C-Level: 0.90, 0.95, 0.99, etc
- Pooled: Select Yes (We are assuming \( \sigma_1 = \sigma_2 \))
Calculate: Select Calculate and press Enter

Program Output:
- Confidence Interval: (lower bound, upper bound)
- Degrees of freedom used for t-distribution
- Mean of first sample
- Mean of second sample
- Standard deviation of first sample
- Standard deviation of second sample
- Pooled sample standard deviation
- Sample size of first sample
- Sample size of second sample

CI for Difference of Population Proportions (\( p_1 - p_2 \))

Select: STAT / TESTS / 2-PropZInt
- \( x_1 \): number of successes found in first sample
- \( n_1 \): sample size of first sample
- \( x_2 \): number of successes found in second sample
- \( n_2 \): sample size of second sample
- C-Level: 0.90, 0.95, 0.99, etc
Calculate: Select Calculate and press Enter

Program Output:
- Confidence Interval: (lower bound, upper bound)
- Calculated value of p-hat statistic of first sample
- Calculated value of p-hat statistic of second sample
- Sample size \( n \) of first sample
- Sample size \( n \) of second sample
Hypothesis Tests Using the TI Graphing Calculator (pages 2 - 4)

Hypothesis Test for Population Proportion $p$

Select: STAT / TESTS / 1-PropZTest
- $p_0$: the population proportion stated in $H_o$
- $x$: number of successes found in sample
- $n$: sample size
- $\text{prop} \neq p_0 < p_0 > p_0$ (select $H_1$ test type)

Calculate: Select Calculate and press Enter
Or
Draw: Select Draw and press Enter

Program Output:
- $H_1$ hypothesis test type
- Value of $z$-standard normal distribution test statistic
- P-value of test statistic
- Calculated value of p-hat statistic
- Size of random sample

Hypothesis Test for Population $\mu$ ($\sigma$ known)

Select: STAT / TESTS / Z-Test
- $\mu_0$: the population $\mu$ stated in $H_o$
- $\sigma$: the standard deviation of the parent pop.
- $\bar{x}$: the sample mean statistic
- $n$: sample size
- $\mu \neq \mu_0 < \mu_0 > \mu_0$ (select $H_1$ test type)

Calculate: Select Calculate and press Enter
Or
Draw: Select Draw and press Enter

Program Output:
- $H_1$ hypothesis test type
- Value of $z$-standard normal distribution test statistic
- P-value of test statistic
- Value of sample mean statistic
- Size of random sample

Hypothesis Test for Population $\mu$ ($\sigma$ unknown)

Select: STAT / TESTS / T-Test
- $\mu_0$: the population $\mu$ stated in $H_o$
- $\bar{x}$: the sample mean statistic
- $S_x$: the sample standard deviation statistic
- $n$: sample size
- $\mu \neq \mu_0 < \mu_0 > \mu_0$ (select $H_1$ test type)

Calculate: Select Calculate and press Enter
Or
Draw: Select Draw and press Enter

Program Output:
- $H_1$ hypothesis test type
- Value of t-distribution test statistic
- P-value of test statistic
- Value of sample mean statistic
- Value of sample standard deviation statistic
- Size of random sample

Test for Difference of Population $p$'s ($p_1 - p_2$)

Select: STAT / TESTS / 2-PropZTest
- $x_1$: the number of successes in first sample
- $n_1$: size of first sample
- $x_2$: the number of successes in the second sample
- $n_2$: sample size of the second sample
- $p_1 \neq p_2 < p_2 > p_2$ (select $H_1$ test type)

Calculate: Select Calculate and press Enter
Or
Draw: Select Draw and press Enter

Program Output:
- $H_1$ hypothesis test type
- Value of $z$-standard normal distribution test statistic
- P-value of test statistic
- Calculated p-hat of first sample
- Calculated p-hat of second sample
- Calculated pooled p-hat statistic of two samples
- Size of first random sample
- Size of second random sample
Test for Difference of Population \( \mu \)'s \( (\mu_1 - \mu_2) \)
Population \( \sigma \)'s are unknown and we do not assume \( \sigma_1 = \sigma_2 \).
Sample statistics are taken from two independent simple random samples. If the parent populations are not normal, the sample sizes should be 30 or more.

Select: STAT / TESTS / 2-SampTTest
Inpt: Use arrow keys – select Stats
x1: sample mean of first sample
Sx1: sample standard deviation of first sample
n1: sample size of first sample
x2: sample mean of second sample
Sx2 : sample standard deviation of second sample
n2: sample size of the second sample
\( \mu_1 \neq \mu_2 \) \( < \mu_2 \) \( > \mu_2 \) (select \( H_1 \) test type )
Pooled: Select No (We are not assuming \( \sigma_1 = \sigma_2 \))
Calculate: Select Calculate and press Enter
Or
Draw: Select Draw and press Enter

Program Output:
\( H_1 \) hypothesis test type
Value of \( t \)-distribution test statistic
P-value of test statistic
Degrees of freedom of the \( t \)-distribution
Sample mean of first sample
Sample mean of second sample
Sample standard deviation of first sample
Sample standard deviation of second sample
Size of first random sample
Size of second random sample

Test for Mean of Paired Differences of Two Dependent Populations - \( d \)
Matched data pairs are taken from two dependent simple random samples of equal size. The random variable \( d \) is a difference statistic where the \( i \)th difference \( d_i = x_i - y_i \) where \( x_i \) is the \( i \)th data value from the first sample and \( y_i \) is the \( i \)th data value from the second sample.

The number of matched pairs should be \( \geq 30 \) or the difference values should come from a normal or almost normal population.

Preliminary Steps:
a) Store the \( x \)-values from the first sample in \( L_1 \).
b) Store the \( y \)-values from the second sample in \( L_2 \).
c) Store the list of differences \( L_1 - L_2 \) in \( L_3 \). Clear the screen and enter the following command:

\[ L_1 - L_2 \rightarrow L_3 \]

(The \( d \)-bar statistic has a \( t \)-distribution.)

Select: STAT / TESTS/ T-Test
Inpt: Use arrow keys – select Data
\( \mu_0 : 0 \) (Ho states that the mean of differences = 0)
List : \( L_3 \)
Freq: 1
\( \mu \neq \mu_0 \) \( < \mu_0 \) \( > \mu_0 \) (select \( H_1 \) test type )
Calculate: Select Calculate and press Enter

Program Output:
\( H_1 \) hypothesis test type
Value of \( t \)-distribution test statistic
P-value of test statistic
Sample mean of the differences in \( L_3 \)
Sample standard deviation of the differences in \( L_3 \)
\( n = \) number of data values in \( L_3 \)
Goodness-of-Fit Hypothesis Test
Test whether or not the observed frequencies of a set of data values have a particular probability distribution. Each data set is partitioned into k categories so that one set contains the observed frequency of each category and the other data set contains the expected frequency of each category.

This test involves calculating the value of a chi-square test statistic and then determine whether or not the value of the \( \chi^2 \) test statistic deviates too far to the right-tail of the \( \chi^2 \) pd curve. Large test statistic values indicate that the observed data frequencies deviate too far from the expected frequencies and therefore the observed frequencies do not fit a particular probability distribution.

Preliminary Steps:
- a) Store the k observed frequencies in L1.
- b) Store the k expected frequencies in L2.

Select: STAT / TESTS / \( \chi^2 \)GOF-Test
Observed : L1
Expected : L2
df : degrees of freedom = number of categories - 1
Calculate: Select Calculate and press Enter

Program Output:
\( \chi^2 = \) value of the chi-square test statistic
P-value of test statistic
df = degrees of freedom

Hypothesis Test for Independence
Test whether or not events in a sample space are independent. The intersection of two events in the sample space corresponds to the intersection of a row and column of a contingency table. Hence the sample space is partitioned into \( r \) disjoint events which correspond to the \( r \) rows of the contingency table. The \( c \) columns of the contingency table correspond to a second partition of the sample space into \( c \) disjoint events. There are two contingency tables; one table contains observed frequencies and the other table contains expected frequencies.

The test involves calculating the value of a \( \chi^2 \) test statistic and then determine whether or not the value of the test statistic deviates too far to the right-tail of the \( \chi^2 \) pd curve. Large right-tail \( \chi^2 \) test values indicate that the observed data frequencies deviate too far from the expected frequencies and therefore the events in the sample space are not independent.

The df parameter of the \( \chi^2 \) pd = \( (r - 1)(c - 1) \) where \( r \) and \( c \) equal the number of rows and columns of a contingency table.

For the test to be valid, the expected frequency of every cell in the expected contingency table must be \( \geq 5 \).

Preliminary Steps:
- Use the TI MATRIX-EDIT menu command to create an \( r \)-row by \( c \)-column matrix of observed frequencies. Enter and save the observed frequencies in matrix [A].
- Use the TI MATRIX-EDIT menu command to create an \( r \)-row by \( c \)-column matrix [B] which will contain the expected frequencies. Do not bother to fill in the cells of [B] since [B] cell values will be automatically calculated and filled in later.

Select: STAT / TESTS / \( \chi^2 \)-Test
Observed : [A]
Expected : [B]
Calculate: Select Calculate and press Enter

Program Output:
\( \chi^2 = \) value of the chi-square test statistic
P-value of test statistic
[B] will now contain the expected frequencies

Hypothesis Test for Homogeneity
Test to see if two or more populations have the same proportions of different characteristics of interest. The test involves two contingency tables and a \( \chi^2 \) test statistic. Each row of a table contains the population frequencies which correspond to the proportions of a population.

The cell values in row \( r \) of one table are the observed population characteristic frequencies of population \( r \). The cell values in row \( r \) of the second table are the expected population characteristic frequencies of population \( r \).

Follow the same procedure for doing a hypothesis test for independence which is described above.
Find Equation of Regression Line \((y = a + bx)\), Sample Correlation Coefficient \(r\) and the Coefficient of Determination \(r^2\) with the TI 83/84+ graphing calculator.

a) Clear lists \(L_1\) and \(L_2\).
b) Enter the x-coordinates in list \(L_1\).
c) Enter the y-coordinates in list \(L_2\).
d) Press \(\text{STAT} / \text{TESTS} / \text{LinRegTTest}\)
   - Xlist : \(L_1\)
   - Ylist : \(L_2\)
   - Freq : 1
   - \(\beta\) and \(\rho\) : \(\neq 0\)
   - RegEQ :
     - Select \(\text{Calculate}\) and press the \(\text{ENTER}\) key.

Program Output:
\[y = a + bx\]
\[\beta \neq 0\] and \(\rho \neq 0\)
\(\beta\) (beta) is a population parameter equal to the true value of the slope of the regression line.
\(\rho\) (rho) is a population parameter equal to the true value of the correlation coefficient.

- \(t\) = value of test statistic derived from a random sample
- \(p\) = P-value of test statistic
- degrees of freedom of t-distribution = \(n - 2\)
- \(a\) = y-intercept of the regression line
- \(b\) = slope of the regression line
- \(s\) = standard error where larger values of \(s\) indicate increased scattering of points
- \(r^2\) = the coefficient of determination
- \(r\) = the sample correlation coefficient

Linear Correlation Hypothesis Test
\[H_0 : \rho = 0\] - There is no linear correlation.
\[H_1 : \rho \neq 0\] - There is linear correlation.

Probability distribution of the test statistic \(t = r / \sqrt{(1 - r^2) / (n - 2)}\) is a t-distribution with \(n-2\) degrees of freedom.

Draw Scatter Plot and Graph Regression Line with TI 83/84+

a) Enter the x-coordinates in list \(L_1\).

b) Enter the y-coordinates in list \(L_2\).

c) Press the \(\text{MODE}\) button.
   - Select \(\text{NORMAL}\) number display mode
   - Select \(\text{FLOAT}\) and set rounding to 4 decimal places
   - Select \(\text{FUNC}\) graph type
   - Select \(\text{CONNECTED}\) plot type
   - Select \(\text{SEQUENTIAL}\)
   - Select \(\text{REAL}\) number mode
   - Select \(\text{FULL}\) screen mode

d) Press the \(\text{WINDOW}\) button. Set axes scale values (Xmin, Xmax, etc.) to fit scatter plot data.

e) Press the \(\text{STAT PLOT}\) key. (\(2\text{ND}\) and \(Y=\))
   - Set Plot 1 to \(\text{on}\) and all other plots to \(\text{off}\)
   - Type : Select scatter plot icon (top-row-left)
   - Xlist : \(L_1\)
   - Ylist : \(L_2\)
   - Mark : Select desired style of plot marker.

f) Press the \(Y=\) button.
   - Clear out all equations with the \(\text{CLEAR}\) key.
   - \(Y_1 = \) equation of regression line: \(bx + a\) or \(a + bx\)

g) Press the \(\text{GRAPH}\) button to view the graph of the regression line.

After the least squares regression line is graphed, points on the regression line can be found as follows: Press the \(\text{CALC}\) (\(2\text{ND TRACE}\)) key and select \(\text{value}\). Then enter a value for the \(x\) variable and press the \(\text{ENTER}\) key. Continue entering other values of \(x\) as desired.
Normal Quantile Plot – Check to see if a sample of \( n \) data points came from a normal population.

Enter the sample data values in list \( L_1 \).

Press the STAT PLOT key. ( 2ND and Y= keys)
- Turn Plot1 on and the other plots to off.
- Type : Select the plot icon in row-2-right.
- Data List : \( L_1 \)
- Data Axis : \( X \)
- Mark : Select the desired data marker style.

Press the ZOOM key and 9 to generate a quantile plot of the sample data values.
Press the TRACE key to view the x-y coordinates of points on the graph.

Warning! Quantile plots of sample data taken from a uniform distribution may appear to be somewhat linear, however, the plot follows a systematic curved pattern about a straight line and therefore it is not considered to be a linear plot.

Example 1: The data set below is a random sample of 16 data values taken from an exponential population with \( \mu = 4 \) and \( \sigma = 4 \). Exponential populations are very skewed to the right and therefore normal quantile plots of samples taken from an exponential population should not be linear.

\[
\{ 0.160, 7.697, 0.552, 2.266, 2.469, 5.254, 8.143, 2.211, 2.346, 3.901, 3.1619, 3.105, 11.821, 0.737, 1.415, 16.282 \}
\]

Example 2: The data set below is a random sample of 21 data values taken from a normal population with \( \mu = 100 \) and \( \sigma = 16 \). Since the sample was taken from a normal population, the normal quantile plot of the data set should be linear or almost linear. See the comment below.

\[
\]

Comment: The \( n \) data values, \( \{ x_i \} \), are first sorted in ascending order. Each \( x_i \) is assigned a percentile rank \( P_i = (i - 0.5) / n \). Each y-coordinate \( y_i = \) the z-score corresponding to \( P_i = \text{invNorm}(P_i, 0, 1) \). You can judge the straightness of a line by eye. Do not pay too much attention to points at the ends of the plot, unless they are quite far from the line. It is common for a few points at either end to stray from the line somewhat. However, a point that is very far from the line when the other points are very close is an outlier, and deserves attention. The Ryan-Joiner test is one of several other tests available to test for normality.
Residual Plot – Plot the residuals of a sample of n (x, y) data points associated with the least-squares regression line of the n data points. The residual of a data point (x,y) equals the difference between y and \( y_p = y - y_p \) where \( y_p \) equals the y-value predicted by the least-squares regression line. A residual equals the unexplained deviation between y and \( y_p \). The least-squares regression line minimizes the sum of the squares of the residuals.

If the residual plot shows a random pattern about the x-axis, the least-squares regression line is a good model to fit the data. If the residual plot pattern about the x-axis is not random, a nonlinear model such as quadratic or cubic model might be a more appropriate model to fit the data.

Select **STAT / SetUpEditor** and press the **ENTER** key.
Clear L₁, L₂, and L₃.
Enter the x-coordinates in L₁.
Enter the y-coordinates in L₂.

Now create a special list for the residuals in the list editor.
- Select **STAT / EDIT**
- Move the edit cursor to the very top of L₃ so that L₃ is highlighted.
- Press the **INS** key. ( 2nd DEL = the insert key)
- Use the command **LIST / NAMES** and select **RESID** from the list and then press the **ENTER** key to paste in the RESID symbol at the top of the new list.

Next calculate the equation of the least-squares regression line and correlation coefficient \( r \).
- Select **STAT / TESTS / LinRegTTest**
- Xlist: L₁
- Ylist: L₂
- Freq: 1
- \( \beta \) and \( \rho : \neq 0 \)
- RegEQ:
- Select **Calculate** and press the **ENTER** key.
- Record the y-intercept \( a \), slope \( b \), \( r \) and \( r^2 \).
- Select **STAT / EDIT** to view the residuals which were automatically inserted in list **RESID**.

Now plot the residuals for each x-y data point.
- Press the **STAT PLOT** key. ( 2ND and Y= keys)
- Turn Plot1 **on** and the other plots to **off**.
- Type: Select the scatter plot icon top-row-left.
- Xlist: L₁
- Ylist: **RESID** ( Select **LIST / NAMES** and then select **RESID** )
- Mark: Select the desired data marker style.
- Select **ZOOM** and 9 to generate a residual plot of the sample x-y data pairs.
- Press the **TRACE** key to view the x-y coordinates of points on the residual plot.

**Comment:** Page 225 of the course textbook has two excellent examples that illustrate how to use a residual plot to determine whether or not a linear model is an appropriate model of the data set.
Construct a Frequency Histogram of a Data Set

- Set the variable \( n \) equal to the number of classes you wish to have.
- Use STAT button to clear list \( L_1 \), enter the data values in list \( L_1 \) and then sort the data values in ascending order.
- Use STAT / EDIT to view the data values in \( L_1 \). Then find \( X_{\text{min}} \) and \( X_{\text{max}} \) and set the class width \( CW = (X_{\text{max}} - X_{\text{min}}) / n \).
  - Always round \( CW \) up to the next whole number even if \( CW \) was a whole number. Example: If \( CW = 9 \), round up to 10.
- Follow the procedure given in a separate handout to calculate class limits, class midpoints and class boundaries.
- Press the STAT PLOT button. (2 ND and \( Y = \) buttons)
  - Select Plot 1.
  - Set Plot 1 to ON.
  - Type : Select the histogram icon in top-row-right. (3 rd icon in list of icons)
  - Xlist: \( L_1 \)
  - Freq: 1
  - Mark : Select desired data graph marker. A dot marker is difficult to see.
- Make sure that all of the other STAT PLOTS are set to OFF.
- Press the Y= button and use the CLEAR button to clear out any equation formulas.
- Press the WINDOW button.
  - Set Xmin = the lower class boundary of the first class = the midpoint of first class - \( CW / 2 \).
  - Set Xmax = the upper class boundary of the last class = the midpoint of the last class + \( CW / 2 \).
  - Set Xscl = the class width = \( CW \).
  - Set Ymin = -1.
  - Set Ymax = a little more than the estimated maximum class frequency. You can change this later if needed.
  - Set Yscl = 2 or 5. Too many tick marks will make the y-axis too crowded.
  - Set Xres = 1 which is the highest possible screen resolution.
- Press the GRAPH button to view the frequency histogram graph.
- Press the TRACE button to view class boundaries and class frequencies. Use the arrow keys to move from class to class.

Construct a Frequency Polygon of a Data Set

- Set the variable \( n \) equal to the number of classes you wish to have.
- Find the class midpoints and class frequencies. Refer to the procedure above for calculating a frequency histogram.
- Let \( Mn = \) the midpoint of the \( n \)th class and \( Fn = \) the frequency of the \( n \)th class.
  - In list \( L_2 \) enter the \( n + 2 \) x-coordinates of the polygon graph : \( M_1, CW, M_2, M_3, \ldots, Mn, Mn + CW \).
  - In list \( L_3 \) enter the \( n + 2 \) y-coordinates of the polygon graph : \( 0, F_1, F_2, F_3, \ldots, Fn, 0 \).
- Press the STAT PLOT button. (2 ND and \( Y = \) buttons)
  - Select Plot 1.
  - Set Plot 1 to ON.
  - Type : Select polygon frequency icon in middle of top-row. (2 nd icon in list of icons)
  - Xlist : \( L_2 \)
  - Ylist : \( L_3 \)
  - Mark : Select desired data graph marker. A dot marker is difficult to see.
- Make sure that all of the other STAT PLOTS are set to OFF.
- Press the Y= button and use the CLEAR button to clear out any equation formulas.
- Press the ZOOM button and 9 to view the frequency polygon graph.
- Press the TRACE button to view class midpoints and class frequencies. Use the arrow keys to move from class to class.

Construct a Box-and-Whisker Plot of a Data Set

- In list \( L_1 \) enter the numbers in the data set.
- Press STAT PLOT button and set Plot 1 to ON and all other plots to OFF.
  - Set the following plot settings: Xlist : \( L_1 \), Type : Select desired box-and-whisker icon, Freq : 1, Mark : Select a graph marker.
- Press the ZOOM button and 9 to view to the graph of the box-and-whisker plot.
- Press the TRACE button to view the key plot points. Use the arrow keys to move from point to point.

One of the box-plot icons is for a plot that does not connect the whiskers to outliers. This makes it easy to identify outliers.