

Minitab Users Guide

to accompany

Statistics: Unlocking the Power of Data

by Lock, Lock, Lock, Lock, and Lock

Getting Started

NOTE: In Minitab, your worksheet is your dataset. Your project is all the analysis, including graphs. Be sure to save your worksheet (as a worksheet not a project) once you have entered your data!

To enter data:

Remember that each column is a different variable and the rows are the cases.

1. If your data already exists in some format, such as Excel, you can import it into Minitab by copying and pasting. Be sure, however, that titles of columns (names of variables) go in the title row only, which is the row directly below C1, C2, etc. Titles can never go below that, where only the data belongs.
2. If you are typing your data in yourself, you enter the values as you would into any spreadsheet, using arrow keys to move about the worksheet. Be sure that column titles (names of variables) go only in the row right below the row with C1, C2, etc. We recommend that you name your variables using something short and descriptive.

Warning: Minitab has quantitative variables, which can contain only numbers, and categorical variables, which can contain anything. *As soon as you enter anything other than a number in a column*, Minitab assumes that the entire column is categorical. This includes things like dollar signs, units, etc. If Minitab has concluded that a column is categorical, you will see a “-T” added to the column name at the top and Minitab will not ever allow you to do quantitative calculations with that column. It will stay a categorical variable forever, even after you delete the offending character.

If you enter something other than a number in a quantitative variable column by mistake, you will have to re-enter the values for that variable in a different column (without any non-numbers!)

Once you have entered or imported your data, be sure to SAVE the data as a Minitab worksheet so that you can use it again.

Using Minitab in Chapter 2

Categorical Variables

Tables for categorical variables:

For a single categorical variable: Starting with the menu across the top, go to

Stat ⇒ Tables ⇒ Tally Individual Variables

With the cursor in the Variables box, double-click on the variable(s) you want.

You can select Counts and/or Percents, then click OK.

For a relationship between categorical variables: Starting with the top menu, go to

Stat ⇒ Tables ⇒ Cross Tabulation and Chi Square

Select a variable for rows and one for columns (you can ignore labels for now).

Select Counts and/or Row Percents and/or Column Percents and/or Total Percents,
Then click OK.

Graphs for categorical variables:

For a single categorical variable: Starting with the menu across the top, go to

Graph ⇒ Bar Chart or Pie Chart

For a relationship between categorical variables: Starting with the top menu, go to

Graph ⇒ Bar Chart, then choose Cluster or Stack

One Quantitative Variable

Statistics for a single quantitative variable:

Starting with the menu across the top, go to

Stat ⇒ Basic Statistics ⇒ Display Descriptive Statistics

Select the variable: With the cursor in the Variables box, double-click on the variable desired
and then click OK.

Graphs for a single quantitative variable:

Starting with the menu across the top, go to

Graph ⇒ Histogram or Dotplot or Boxplot

Choose “Simple” for now. Select the variable as above and click OK.

Note: If you want your boxplots to be horizontal, click on the “scale” button and then click on
the box to transpose scales. Then click OK.

Using Minitab in Chapter 2, continued

One Quantitative Variable by categories in One Categorical Variable

Statistics for a quantitative variable by categories in a categorical variable:

Starting with the menu across the top, go to

Stat ⇒ Basic Statistics ⇒ Display Descriptive Statistics

Select the quantitative variable by putting the cursor in the “Variables” box and double-clicking on the variable desired, then select the categorical variable by putting the cursor in the “By” box and double-clicking on the categorical variable desired.

Graphs for a quantitative variable by categories in a categorical variable:

Starting with the menu across the top, go to

Graph ⇒ Boxplot (or try histograms or dotplots)

Choose the option in the upper right: “One Y: With Groups”. Select the quantitative variable for the “Graph variable” by putting the cursor in that box and double-clicking on the variable desired. Select the categorical variable by putting the cursor in the “Categorical variable” box and double-clicking on the variable desired.

Note: If you want your boxplots to be horizontal, click on the “scale” button and then click on the box to transpose scales. Then click OK.

Two Quantitative Variables

Statistics for two quantitative variables:

Correlation: Starting with the menu across the top, go to

Stat ⇒ Basic Statistics ⇒ Correlation

Select the two quantitative variables. (If you select more than two, you will get a correlation matrix.)

Linear Regression: Starting with the menu across the top, go to

Stat ⇒ Regression ⇒ Regression

Select the response variable and then the explanatory variable, and click OK.

Graphs for two quantitative variables:

Starting with the menu across the top, go to

Graph ⇒ Scatterplot

Choose either “Simple” (without the regression line) or “With Regression” if you want the regression line added. Enter the Y (response) variable and X (explanatory) variable. Then click OK.

Using Minitab for Chapters 3 and 4

The current version of Minitab (16.0) has no procedures for creating bootstrap or randomization distributions. We believe the StatKey tools at lock5stat.com/statkey provide better options for these procedures.

However, for those instructors wishing to do so, we have produced a set of Minitab macros which instructors can download from the instructor resources site. These macros enable students to create bootstrap and randomization distributions from within Minitab. These are text files with a .mac extension that execute a series of Minitab commands. They are a bit clunky but available for those who want them. Details are below.

To execute a macro from within Minitab:

- From the Session Window, select *Editor > Enable Commands* to show the Minitab command prompt.
- Run a macro by typing %macroname along with any extra information needed (such as a column of input data or column to contain the simulated statistics).
- Minitab will look for the macro file in its default macro folder or the current data folder (e.g. put a copy of the macro wherever you load the data worksheet). You can also specify a path as part of the macro name if it is located elsewhere.

Example: Suppose we want to construct a distribution of bootstrap means for the prices of used Mustangs based on the data in **MustangPrice.mtw**. Load the worksheet into Minitab and note that the *Price* variable is stored in C3 and column C4 is empty so it can contain the bootstrap means. The macro for bootstrapping a single mean (or median or std. dev.) is called **OneMeanBoot**.

```
MTB > %OneMeanBoot 'Price' c4
Executing from file: onemeanboot.MAC

Which statistic?
1=mean    2=median    3=std dev
DATA> 1

How many bootstrap samples?
DATA> 1000
```

This creates new data in C4 (and names in BootStat) with the 1000 bootstrap means. Be careful that you don't already have important data in the column you choose to hold the simulated statistics!

Here are the available macros:

OneMeanBoot C3 C4	C3=data, C4= hold bootstrap statistics
OneMeanRand C3 C4	C3=data, C4= hold randomization statistics
OnePBoot C4	C4= hold the bootstrap proportions'
OnePRand C4	C4= hold the randomization proportions
TwoMeanBoot C2 C3 C4	C2=quantitative C3=groups C4= hold bootstrap differences
TwoMeanRand C2 C3 C4	C2=quantitative C3=groups C4= hold bootstrap differences
SlopeBoot C2 C3 C3	C2=response C3=predictor C4= hold bootstrap slopes (or r)
SlopeRand C2 C3 C4	C2=response C3=predictor C4= hold randomization slopes (or r)

Once the column of bootstrap/randomization statistics is created, you can use the usual plot and descriptive statistics commands to work with it. Also, you can **Data>Sort** the column to count number of values beyond a threshold or use Calc>Calculator and a function like **percentile (C5, 0.025)** to find a percentile.

Using Minitab for Theoretical Distributions in Chapters 5 – 10

Finding Values for Theoretical Distributions

To find a probability or a percentile from a theoretical distribution, go to:

Graph => Probability Distribution Plot => Select “View Probability” and click OK.

You will see the “Distribution” screen, and can select Normal, t, or other distributions off the drop-down menu, and enter any relevant parameters (such as df for the t-distribution).

You should then select the “Shaded Area” tab. You can opt to put in a Probability (area) or an X-value (cutoff or boundary value), and you can select Left tail, Right tail, or Two-tail or Middle. Make the selections and click OK and you will see the result.

Using Minitab in Chapter 6

Inference for Means: t-Intervals and t-Tests

Confidence Interval for a Mean:

Select **Stat => Basic Statistics => 1-Sample t.**

Click in the box for “Samples in Column” and then select the variable. If you then click OK, you will get a 95% confidence interval. If you want a different confidence level, select “Options” and change the level.

Hypothesis Test for a Mean:

Select **Stat => Basic Statistics => 1-Sample t.**

Click in the box for “Samples in Column” and then select the variable. Check “Perform Hypothesis Test” and then enter the null hypothesis value. If you want a one-tailed alternative, select “Options” and change the alternative.

Hypothesis Test and Confidence Interval for a Difference in Means:

Select **Stat => Basic Statistics => 2-Sample t.**

Click in the box for “Samples in One Column” and then select the relevant quantitative variable for “Samples” and the categorical variable for “Subscripts”. (Note that the categorical variable can only have two possible categories for a difference in means test.) If you want a different confidence level for the interval or if you want a one-tailed alternative hypothesis test, select “Options”.

Inference for Proportions: Intervals and Tests

Confidence Interval for a Proportion:

Select **Stat => Basic Statistics => 1-Proportion.**

Click in the box for “Samples in Column” and then select the variable. If you then click OK, you will get a 95% confidence interval. If you want a different confidence level, select “Options” and change the level.

Hypothesis Test for a Proportion:

Select **Stat => Basic Statistics => 1-Proportion.**

Click in the box for “Samples in Column” and then select the variable. Check “Perform Hypothesis Test” and then enter the null hypothesis value. If you want a one-tailed alternative, select “Options” and change the alternative.

Hypothesis Test and Confidence Interval for a Difference in Proportions:

Select **Stat => Basic Statistics => 2-Proportions.**

Click in the box for “Samples in One Column” and then select the relevant categorical variable for “Samples” and the relevant categorical variable for “Subscripts”. (Note that the categorical variables can only have two possible categories for a difference in proportions test.) If you want a different confidence level for the interval or if you want a one-tailed alternative hypothesis test, select “Options”.

Using Minitab for Chapter 7

Chi-Square Tests

Chi-Square Goodness-of-Fit Test for a Single Categorical Variable

Select **Stat => Tables => Chi-Square Goodness-of-Fit Test**

If your categorical data is in a column of a dataset, select “Categorical Data” and click in the box and select your categorical variable. It is also possible to put in counts – if you choose to do this, use “Observed Counts” and enter the observed counts, separated by spaces.

If you are testing that the categories are equally likely, leave the “Equal Proportions” option checked. If, instead, you have different assumed proportions for the different categories, select “Specific Proportions” and change the “Input column” box to “Input constants”. A window will open allowing you to enter the null hypothesis assumed proportions. Then click OK.

The test gives the observed counts, the expected counts, and the contribution to the chi-square statistic, as well as the final chi-square test statistic and the p-value. It also opens several graphs related to the test.

Chi-Square Test for Association for Two Categorical Variables

Select **Stat => Tables => Cross Tabulation and Chi-Square**

If your categorical variables are in two columns of a dataset, enter them in the boxes “For rows” and “For columns”. (It doesn’t matter which is which.) It is also possible to enter the counts in a two-way table directly – see below for details.

Click on the Chi-Square button and then select “Chi-Square analysis”. (We recommend also requesting “Expected cell counts” and “Each cell’s contribution to the Chi-Square statistic”.) Click OK and then OK again.

The test then gives the observed counts, the expected counts, and the contribution to the chi-square statistic, as well as the final chi-square test statistic and the p-value.

To enter counts from a two-way table directly:

Enter the counts from the two-way table (not including totals) directly into a blank worksheet and then:

Select **Stat => Tables => Chi-Square Test (Table in Worksheet)**

Enter the columns containing the two-way table counts into the box provided and then click OK.

Using Minitab for Chapter 8

Analysis of Variance for Difference in Means

Select **Stat => ANOVA => One-Way Analysis of Variance**

Enter the quantitative variable as the “Response” and the categorical groups variable as the “Factor” then click OK. You will see the analysis of variance table, including the p-value, as well as some summary statistics for each of the groups. Depending on the version of Minitab you have, you might also see 95% confidence intervals for each of the group means based on the pooled standard deviation, as described in Section 8.2.

Using Minitab for Chapters 9 and 10

Correlation

Select **Stat => Basic Statistics => Correlation**

If you enter two quantitative variables in the box and click OK, you will see the correlation between the two variables and the p-value for a two-tailed test of that correlation. (Make sure “Display p-values” is checked in order to see the p-value.)

If you enter more than two quantitative variables in the box and then click OK, you will see a correlation matrix. In this matrix, the rows and columns are labeled with the variables, and the correlation (with p-value below it) is shown in the matrix entry for that row and that column. Again, make sure “Display p-values” is checked in order to see the p-values.

Linear Regression

Select **Stat => Regression => Regression**

Enter the Response variable in the “Response” box and the explanatory variable(s) in the “Predictors” box, and click OK. You will see:

- The regression equation
- The standard error, test statistic (T), and p-value (P) for the slope(s)
- The value of R-squared
- The ANOVA for regression table, including the p-value for the ANOVA test
- Miscellaneous other pieces of information

See Chapters 9 and 10 for an explanation of all output.