PHStat2 Help

version 3.52

Table of Contents

1 Welcome to PHStat2

1.1 Welcome

Welcome to PHStat2, the software that you use with Microsoft Excel to explore statistics.

If you have used PHStat2 previously, click What's New? to learn about the changes in this version of PHStat2.

If you are a new user of PHStat2, read the "PHStat2 Quick Start Guide" topics, beginning with What is PHStat2?.

You should also review the PHStat2 readme file for any late-breaking news about PHStat2.

(This is help system version 3.52)

1.2 What's New?

PHStat2 version 3.5 is the first PHStat2 version to use the newer Excel function names when it runs in Excel 2010, 2011, or 2013. Version 3.5 does not require the Microsoft-supplied Data Analysis ToolPak and Analysis ToolPak–VBA add-ins.

This version also contains a number of enhancements and improvements over previous PHStat2 versions. Click a topic for more information.

New Statistical Procedures

PHStat2 Revision History

1.2.1 New Statistical Procedures

Descriptive Summary

1.2.2 PHStat2 Revision History

Changes in version 3.50 since version 3.21:

Procedures use newer Excel function names if PHStat is used with Excel 2010, 2011, or 2013.

Corrected issues with the Paired t Test lower-tail option, the Stepwise Regression backwards elimination option, the One-way Tables & Chart Percentage column option, and Simple Linear Regression ANOVA table.

Enhanced the formatting and display of results for selected procedures.

Corrected issues associated with 2013 OS X Excel 2011 updates.

Changes in version 3.21 since version 3.06:

Added OS X Excel 2011 and Microsoft Windows Excel 2013 compatibility.

Added compatibility with 64-bit versions of Microsoft Windows Excels.

Descriptive Summary procedure added.

One-Way ANOVA and Levene Test procedures accept groups of unequal sample sizes.

Corrected an issue in Levene Test procedure when data contains no labels.

Corrected an issue in the Separate-Variance t procedure when the hypothesized difference is not zero.

Corrected an issue in the Two-Tail Test option of the Chi-Square Test for the Variance procedure.

Simplified setup, no special setup.exe program needed to use PHStat; PDF help manual.

Eliminated full-support for Microsoft Windows versions of Excel older than Excel 2007.

Changes in version 3.06 since version 3.0:

Enhanced introductory help system topics.

Corrected minor issues in the Separate-Variance t procedure.

Added notes to help entries for Stepwise Regression, Levene, and One-Way ANOVA procedures.

Updated formatting of worksheets to match Excel 2010 styling. Corrected several cases in which symbols in labels displayed as roman letters.

Corrected minor issues in the Two-way ANOVA with replication and Levene Test procedures.

Corrected formatting of the Variance Inflationary Factor option of Multiple Regression.

Corrected issue with Paired t Test procedure when Upper-Tail Test option is selected.

2 PHStat2 Quick Start Guide

2.1 What is PHStat2?

PHStat2 is software that makes operating Microsoft Excel as distraction-free as possible. As a student studying statistics, you can focus solely on mastering statistics now and not worry about having to become an expert user of Excel at the same time.

PHStat2 executes for you the low-level menu selection and worksheet entry tasks that are associated with implementing statistical analysis in Microsoft Excel. Unlike other add-ins that report results as a series of text labels, hiding the details of using Microsoft Excel, PHStat2 creates actual working Excel worksheets and chart sheets from which you can learn real Excel techniques when you are ready to do so. This makes PHStat2 particularly useful if you are concerned that using a textbook add-in now will leave you unprepared later when you use Microsoft Excel outside your statistics course.

You use PHStat2 to perform a particular analysis by first selecting choices from the PHStat2 menu and completing entries in a dialog box. PHStat2 generates new worksheets and chart sheets to hold the results of an analysis. Many of these sheets contain formulas that will allow you to change the underlying data and see new results without having to rerun a procedure.

This version of PHStat2 works best with Excel 2010 or Excel 2013 (Microsoft Windows) or Excel 2011 (OS X). This version is also compatible with Excel 2007. (This version has limited compatibility with Excel 2003.)

2.2 What PHStat2 is not

PHStat2 is not designed as a commercial statistical package and should not be used as a replacement for packages such as SAS or SPSS. Because PHStat2 emphasizes learning, PHStat2 tries to use the methods of calculation that a learner could easily follow and would find presented in a

textbook that discusses manual calculation. In other cases, PHStat2 uses pre-existing Microsoft Excel methods to produce results that are too complex to manually calculate or present to the learner. Both of these choices will occasionally cause PHStat/Excel to produce results not as precise as those produced by commercial packages that use methods of calculation fine-tuned to the limitations of computing technology. While these differences are not significant for most sets of data, real-world data sets with unusual numerical properties or with extreme values could produce significant anomalies when used with PHStat2 and Microsoft Excel.

If you are a student, use PHStat2 only with the data discussed in your text or required by your instructor. For such sets of data, the results produced by PHStat2 will allow you to form the proper conclusion about your data. (See also How PHStat2 presents its results.)

2.3   How PHStat2 Presents Its Results

PHStat2 presents its results in new worksheets and chart sheets inserted into the currently active workbook. If the worksheet(s) produced are interactive, PHStat2 tints user changeable cells light turquoise and tints the cells containing the results in light yellow. The yellow-tinted results cells and other cells that hold intermediate calculations are usually only minimally formatted to allow you to see the value that Microsoft Excel is actually uses when computing the formulas in the worksheet. For some procedures, these values will have an excessive number of seemingly significant digits and you may want to rework or reformat them for later presentation.

2.4   Preparing Data for Analysis

Prepare your data for analysis by PHStat by placing it in columns on a new worksheet beginning with column A and row 1. If you are selecting only part of a column, copy that partial column range to a new column or worksheet. This good practice will allow you to easily verify your data later as well as avoid some common types of user errors.

Make the worksheet that contains the data to be analyzed the currently active worksheet before you select a PHStat procedure. (Clicking a worksheet's sheet tab, the operation you do to view the contents of a sheet, makes a worksheet the currently active sheet.) If you forget to make the data worksheet the currently active worksheet, you may encounter an error message that refers to an unexpected error or an error in handling a cell range.

For procedures that require two or more cell ranges, such as the regression procedures, make sure that all of your cell ranges are from the same worksheet.

Generally, you should use row 1 to enter a label for a column's data. Due to technical limitations of Excel, avoid using numbers as a row 1 labels. If you must enter a number, enter the number preceded by an apostrophe. For example enter '2014.

Save your workbook using the **.xlsx** format, not the older **.xls** format.

Sometimes opening an Excel workbook that was created in an earlier version of Excel will cause Excel to go into "compatibility mode." If you see "Compatibility Mode" in the title bar of Excel, save your workbook using the **.xlsx** format, *close the workbook*, and then *reopen* the workbook before using PHStat2. That will ensure that both Excel and PHStat2 will work as intended.

2.5   Using PHStat2

You use PHStat2 by opening the PHStat2.xlam file from a folder or desktop. If you prefer, you can use the Excel File Open dialog box to open this file.

Depending on your security settings, Excel may display a dialog box that warns about the possibility of macro viruses when PHStat2 opens in Excel. Should this dialog box appear, click the Enable Macros button to permit PHStat2 to be loaded. Users of Microsoft Windows versions of Excel may also need to adjust other security settings, as detailed in the PHStat2 Technical Reference section of this help system.

(See also the sections Getting Started with PHStat and Updating PHStat).

3   PHStat2 Technical Reference

3.1   PHStat2 Technical Requirements

Microsoft Windows Excel 2007, Excel 2010, or Excel 2013, OS X Excel 2011, or Office 365. Limited capability with Microsoft Windows Excel 2003 as explained the Readme file.

For Microsoft Windows Excel versions, Trust Center Macro Settings set to **Disable all macros with notification** (recommended) or **Enable all macros** (not recommended). See the "Configuring Microsoft Windows Excel Security" topic for more details.

Internet access for downloading Microsoft Excel updates from Microsoft websites as they become available.

3.2   Configuring Microsoft Windows Excel Security

To use PHStat2 successfully on Microsoft Windows versions, you first must review and change, if necessary, Microsoft Office security settings. Instructions differ from version to version. This step is not needed if using OS X Excel 2011.

Excel 2010 and Excel 2013 Security

In Excel 2010 or Excel 2013, you must change the Trust Center settings to allow PHStat to properly function. To do so, click **File ➜ Options** and in the Excel Options dialog box:

1. Click **Trust Center** in the left pane and then click **Trust Center Settings** in the right pane.

In the Trust Center dialog box:

2. Click **Add-ins** in the next left pane, and in the Add-ins right pane clear all of the check boxes.

3. Click **Macro Settings** in the left pane and in the Macro Settings right pane click **Disable all macros with notification**.

4. Click **OK** to close the Trust Center dialog box.

Back in the Excel Options dialog box:

5. Click **OK** to finish

For systems with very stringent security settings, you may need to modify step 5. For such systems, in step 5, click **Trusted Locations** in the left pane and then, in the Trusted Locations right pane, click **Add new location**, click **Browse** and navigate to the folder into which you place the PHStat files. Then

click **OK** in the open dialog boxes until all are closed.

Excel 2007 Security

You must change the Trust Center settings to allow PHStat to properly function. Click the **Office Button** and then click **Excel Options** (at bottom of the Office Button menu window). Then continue with steps 1 through 5 of the Excel 2010 and Excel 2013 Security instructions.

Excel 2003 Security

The Microsoft Office macro security level must be set to **Medium** in order to allow PHStat to properly function. Open Microsoft Excel and select **Tools ➔ Macro ➔ Security**. In the **Security Level** tab of the Security dialog box that appears, click the **Medium** option and then click **OK**.

Setting the security level in this dialog box, affects all Microsoft Office programs. When you are finished using PHStat, you can set the security level to **High**, for greatest security. (If you set the level to High, remember to reset the level to Medium before you next open and use PHStat.).

If you cannot see the Macro choice on the Tools menu, it may be hidden from view due to the way that Microsoft Office is currently displaying menu choices. To double-check this, open Microsoft Excel and select **Tools ➔ Customize**. In the Customize dialog box, clear (uncheck) the **Menus show recently used commands first** check box if it is checked and click **Close**.

3.3   Updating PHStat2

From time to time, PHStat2 may be updated to correct minor issues or to add new capabilities.

Visit the web page, MyStatLab course page, or institutional site from which you initially downloaded your copy of PHStat2 to see if an updated version is available. Downloading the PHStat readme file from these sources and comparing its exact decimal version number to 3.52 (this version) is an easiest way to determine if a newer version exists. Depending on the textbook you are using and the licensing that governs your use of PHStat2, you may be able to download a newer version, if it exists.

4   Data Preparation Procedures

4.1   Stack Data

Takes group data that has been arranged in columns and stacks it into two new columns, the first one of which contains the group labels, on a new worksheet.

Data:

Unstacked Data Cell Range: The multiple-column cell range containing the data, arranged one column per group. Must be consecutive columns.

First cells contain group labels: If selected, the contents of the first cell in each column is treated as a group label. If not selected, all cells in the column are treated as data and group labels in the form *Group1*, *Group2*… are created.

4.2   Unstack Data

Takes grouped data that has been placed in two columns, one of which contains group labels, and unstacks it into a series of columns, one for each group, on a new worksheet.

Data:

Grouping Variable Cell Range: The single column cell range containing the group labels.

Stacked Data Cell Range: The single-column cell range containing the data. Must be found on the same worksheet as the grouping variable cell range.

First cells in both ranges contain label: If selected, the contents of the first cell in each column is treated as a descriptive label and not as a data value to be included in the unstacked data.

5   Descriptive Statistics Procedures

5.1   Boxplot

Creates boxplots on a new chart sheet from one or more groups of numerical data. Optionally generates a five-number summary on a separate worksheet.

Data:

Raw Data Cell Range: The single- or multiple-column cell range (see Input Options) containing the numerical data to be plotted.

First cell contains label: If selected, the contents of the first cell in each column of the raw data cell range is treated as a group label. If not selected, all cells in the column are treated as data and group labels in the form *Group1*, *Group2*… are created.

Input Options:

Single Group Variable: If selected, the raw data are interpreted as members of one group. When this option is selected, the raw data cell range must be a single column.

Multiple Groups – Unstacked: If selected, the raw data cell range is interpreted as an unstacked data range in which each column of the range contains the data for an individual group. When this option is selected, the raw data cell range must be two or more consecutive columns.

Multiple Groups – Stacked: If selected, the raw data cell range is interpreted as a stacked series of data values. When this option is selected, the raw data cell range must be a single column.

Grouping Variable Range: The single-column cell range containing the group labels. (Enabled only if Multiple Groups – Stacked is selected.)

Output Options:

Title: The custom title to be placed on the chart sheet and optional worksheet.

Five-Number Summary: If selected, generates a five-number summary table on a new worksheet.

Notes:

This procedure also generates a hidden worksheet that contains the data necessary to generate the boxplot(s).

If the multiple groups – unstacked option is used with a cell range containing stacked data, an error

will occur.

If the multiple groups – stacked option is used with a cell range containing unstacked data, an error will occur.

5.2   Descriptive Summary

Creates a table of descriptive statistics.

Data:

> **Raw Data Cell Range:** The single- or multiple-column cell range (see Input Options) containing the numerical data to be summarized.

> **First cell contains label:** If selected, the contents of the first cell in each column of the cell range entered in the dialog box is treated as a descriptive label and not as a data value.

Input Options:

> **Single Group Variable:** If selected, the variable cell range is interpreted as data for a single group. When this option is selected, the variable cell range should be a single column. If a multiple-column cell range is entered, only the data from the first column of the range is used.

> **Multiple Groups – Unstacked:** If selected, the variable data cell range is interpreted as an unstacked data range in which each column of the range contains the data for an individual group. When this option is selected, the variable data cell range must be two or more consecutive columns.

> **Multiple Groups – Stacked:** If selected, the variable cell range is interpreted as a stacked series of data values. When this option is selected, the variable cell range must be a single column.

>> **Grouping Variable Cell Range:** The single-column cell range containing the group labels. (Enabled only if Multiple Groups – Stacked is selected.

Output Options:

> **Title:** The custom title for the worksheet generated by this procedure. For multiple group data, a group label is appended to this value.

Notes:

If the multiple groups – unstacked option is used with a variable cell range containing stacked data, an error will occur.

If the multiple groups – stacked option is used with a variable cell range containing unstacked data, an error will occur.

5.3   Dot Scale Diagram

Creates a dot scale diagram as a chart object on a new worksheet that also contains the numerical data plotted and a summary table of statistics.

Data:

> **Variable Cell Range:** The single-column cell range containing the numerical data to be plotted.

> **First cell contains label:** If selected, the contents of the first cell in the variable cell range is treated as a variable label and not as a data value to be included in the diagram.

5.4   Frequency Distribution

Creates a table that contains frequency counts and percentage frequency values on a new worksheet. When used with a multiple group option, the frequency table for each group is placed on a separate worksheet. (This procedure enhances and corrects the frequency tables generated by the Data Analysis Histogram tool.)

Data:

> **Variable Cell Range:** The single- or multiple-column cell range (see Input Options) containing the numerical data to be summarized.

> **Bins Cell Range:** The single column cell range containing an ordered list of class maximum values in ascending order that is used to define the bins (classes). The last value in this cell range should be greater than or equal to all values contained in the variable cell range.

> **First cell in each range contains label:** If selected, the contents of the first cell in each cell range entered in the dialog box is treated as a descriptive label and not as a data value.

Input Options:

> **Single Group Variable:** If selected, the variable cell range is interpreted as data for a single group. When this option is selected, the variable cell range should be a single column. If a multiple-column cell range is entered, only the data from the first column of the range is used.

> **Multiple Groups – Unstacked:** If selected, the variable data cell range is interpreted as an unstacked data range in which each column of the range contains the data for an individual group. When this option is selected, the variable data cell range must be two or more consecutive columns.

> **Multiple Groups – Stacked:** If selected, the variable cell range is interpreted as a stacked series of data values. When this option is selected, the variable cell range must be a single column.

>> **Grouping Variable Cell Range:** The single-column cell range containing the group labels. (Enabled only if Multiple Groups – Stacked is selected.)

Output Options:

> **Title:** The custom title for the worksheet generated by this procedure. For multiple group data, a group label is appended to this value.

Notes:

If the multiple groups – unstacked option is used with a cell range containing stacked data, an error will occur.

If the multiple groups – stacked option is used with a cell range containing unstacked data, an error will occur.

If there are values contained in the variable cell range that are greater than the last value in bins cell range, the procedure includes those values in the counts of the last bin (class) and appends a plus sign to the bin label.

5.5   Histogram & Polygons

Creates a frequency table and a histogram (as a chart object) on a new worksheet and optionally generates polygons on separate chart sheets. When used with a multiple group option, the frequency table and histogram for each group is placed on a separate worksheet and the multiple group data are plotted together on the same polygons. (This procedure enhances and corrects the frequency tables

and histograms generated by the Data Analysis Histogram tool.)

Data:

Variable Cell Range: The single- or multiple-column cell range (see Input Options) containing the numerical data to be summarized.

Bins Cell Range: The single column cell range containing an ordered list of class maximum values in ascending order that is used to define the bins (classes). The last value in this cell range should be greater than or equal to all values contained in the variable cell range.

Midpoints Cell Range: Single column cell range containing an ordered list of class midpoints in ascending order. Mandatory if either frequency polygon or percentage polygon output option is selected; optional for all other cases.

First cell in each range contains label: If selected, the contents of the first cell in each column of the cell range entered in the dialog box is treated as a descriptive label and not as a data value.

Input Options:

Single Group Variable: If selected, the variable cell range is interpreted as data for a single group. When this option is selected, the variable cell range should be a single column. If a multiple-column cell range is entered, only the data from the first column of the range is used.

Multiple Groups – Unstacked: If selected, the variable data cell range is interpreted as an unstacked data range in which each column of the range contains the data for an individual group. When this option is selected, the variable data cell range must be two or more consecutive columns.

Multiple Groups – Stacked: If selected, the variable cell range is interpreted as a stacked series of data values. When this option is selected, the variable cell range must be a single column.

Grouping Variable Cell Range: The single-column cell range containing the group labels. (Enabled only if Multiple Groups – Stacked is selected.

Output Options:

Title: The custom title for the worksheet generated by this procedure. For multiple group data, a group label is appended to this value.

Histogram: If selected, generates a histogram (as a chart object) on the worksheet containing the frequency table. Selected by default.

Frequency Polygon: If selected, generates a frequency polygon on a separate chart sheet.

Percentage Polygon: If selected, generates a percentage polygon on a separate chart sheet.

Cumulative Percentage Polygon (Ogive): If selected, generates a cumulative frequency polygon (ogive) on a separate chart sheet.

Notes:

If the multiple groups – unstacked option is used with a variable cell range containing stacked data, an error will occur.

If the multiple groups – stacked option is used with a variable cell range containing unstacked data, an error will occur.

If there are values contained in the variable cell range that are greater than the last value in bins cell range, the procedure includes those values in the counts of the last bin (class) and appends a plus sign to the bin label.

5.6   Scatter Plot

Creates a scatter plot.

Data:

Y Variable Cell Range: The single-column cell range containing the values of the Y variable.

X Variable Cell Range: The single-column cell range containing the values of the X variable. Range must be on the same worksheet as the Y variable cell range.

First cells in both ranges contain label: If selected, the contents of the first cell in each cell range treated as a descriptive label and not as a data value

Output Options:

Title: The custom title for the worksheet generated by this procedure.

5.7   Stem-and-Leaf Display

Creates a stem-and-leaf display using the rounding method on a new worksheet from a set of numerical data. Optionally generates a summary statistics table on the same worksheet.

Data:

Variable Cell Range: The single-column cell range containing the numerical data to be summarized.

First cell contains label: If selected, the contents of the first cell of the variable cell range is treated as a descriptive label and not as a data value.

Stem Unit:

Autocalculate stem unit: If selected, the procedure selects the power of ten which, when divided into the data values, produces at least 5 stems. This is the default option.

Set stem unit as: If selected, the edit box to the right of this option is enabled and a power of ten can be entered to override the value that the procedure would otherwise use. Use this option if the Autocalculate option produces too few or too many stems.

Output Options:

Title: The custom title for the worksheet generated by this procedure.

Summary Statistics: If selected, generates a summary statistics table on the worksheet containing the stem-and-leaf plot.

Notes:

If a value that is not a power of ten is entered as the set stem unit as value, an error will occur.

Use the Set stem unit as option with care as a poor choice of stem unit value could lead to a very many (or very few) stems.

The procedure may produce a plot containing inaccurate data, if the range of the data set is very large and many significant digits are needed to convey the largest and smallest value, for example, if the largest value was 1000003 and the smallest value was .0000001.

5.8   One-Way Tables & Charts

Creates a one-way summary table (in the form of a Microsoft Excel PivotTable) on a new worksheet for a set of categorical data. Optionally generates a bar chart, a pie chart, and a Pareto diagram on separate chart sheets.

Data:

Type of Data:

Raw Categorical Data: If selected, the single-column cell range entered in the dialog box contains raw, unsummarized data. Selected by default.

Table of Frequencies: If selected, the two-column cell range entered in the dialog box contains a frequency table containing categories and frequency counts for each category.

Raw Data Cell Range/Freq. Table Cell Range: The single-column (raw data) or two-column (freq. table) cell range containing the categorical data to be summarized. The label for this entry reflects the type of data option selected.

First cell contains label/First row of table contains label: If selected, the contents of the first cell (or first row) is treated as a descriptive label and not as a data value. The label for this entry reflects the type of data option selected.

Output Options:

Title: The custom title to be placed on the worksheet and any chart sheets generated.

Percentage Column: If selected, adds a column of percentages to summary table.

Bar Chart: If selected, generates a bar chart on a separate chart sheet.

Pie Chart: If selected, generates a pie chart on a separate chart sheet.

Pareto Chart: If selected, adds to the summary table additional frequency, percentage, and cumulative percentage columns sorted in descending order by frequency count and generates a Pareto chart on a separate chart sheet.

Notes:

If the cell range for table of frequencies includes row or column totals, an error will occur.

5.9 Two-Way Tables & Charts

Creates a two-way summary table (in the form of a Microsoft Excel PivotTable) on a new worksheet for bivariate categorical data. Optionally generates a side-by-side bar chart on a separate chart sheet.

Data:

Row Variable Cell Range: The single-column cell range containing the categories to be used as rows in the summary table.

Column Variable Cell Range: The single-column cell range containing the categories to be used as columns in the summary table.

First cell in each range contains label: If selected, the contents of the first cell is treated as a descriptive label and not as a category value.

Output Options:

Title: The custom title to be placed on the worksheet and any chart sheets generated.

Side-by-Side Bar Chart: If selected, generates a side-by-side bar chart on a separate chart sheet.

6   Decision-Making Procedures

6.1   Covariance and Portfolio Management

Creates a new worksheet that calculates the covariance of two variables. The worksheet generated contains an empty probabilities & outcomes table into which you enter probability and outcome data for each variable to complete the analysis. Optionally generates a portfolio management analysis of two investments using a weight-assigned-to-X value that you can change interactively in the worksheet.

Data:

Number of Outcomes: The number of possible outcomes for each variable.

Output Options:

Title: The custom title for the worksheet generated by this procedure.

Portfolio Management Analysis: If selected, adds rows to the new worksheet that perform an analysis of two investments.

6.2   Expected Monetary Value

Creates a new worksheet that calculates expected monetary value from probability and payoff data for the events associated with a set of alternative courses of action. The worksheet generated contains an empty probabilities & payoffs table into which you enter probability and payoff data for each combination of event and alternative course of action to complete the analysis. (You can also enter labels for each event and each alternative course of action.) Optionally generates measures of variation for each alternative course of action and an expected opportunity loss analysis.

Data:

Number of Events: The number of events for each alternative course of action.

Number of Alternative Actions: The number of alternative courses of action, or choices, for a decision.

Output Options:

Title: The custom title for the worksheet generated by this procedure.

Expected Opportunity Loss: If selected, adds rows to the new worksheet that calculate the expected opportunity loss.

Measures of Variation: If selected, adds rows to the new worksheet to calculate measures of variation for each alternative course of action.

Notes:

The sum of all probabilities entered must equal 1. If the sum of all probabilities is another number, the results shown on the worksheet will be invalid.

Worksheets that calculate the expected opportunity loss and/or the measures of variation will initially contain cells that display the Microsoft Excel #N/A and #DIV/0! messages. These "error" messages will disappear when you completely fill in the probability & payoffs table.

Entering the same payoff value for all of the events associated with an alternative course of action (an error) will cause the #DIV/0! message to be displayed in the Return to Risk Ratio cell for that action.

6.3   Expected Opportunity Loss

Creates a new worksheet that calculates expected opportunity losses from probability and opportunity loss data for the events associated with a set of alternative courses of action. The worksheet generated contains an empty probabilities & opportunity losses table into which you enter probability and opportunity loss data for each combination of event and alternative course of action to complete the analysis. (You can also enter labels for each event and each alternative course of action.)

Data:

Number of Events: The number of events for each alternative course of action.

Number of Alternative Actions: The number of alternative courses of action, or choices, for a decision.

Output Options:

Title: The custom title for the worksheet generated by this procedure.

Notes:

The sum of all probabilities entered must equal 1. If the sum of all probabilities is another number, the results shown on the worksheet will be invalid.

6.4   Opportunity Loss

Creates a new worksheet that calculates opportunity losses from payoff data for the events associated with a set of alternative courses of action. The worksheet generated contains an empty payoff table into which you enter payoff data for each combination of event and alternative course of action to complete the analysis. (You can also enter labels for each event and each alternative course of action.)

Data:

Number of Events: The number of events for each alternative course of action.

Number of Alternative Actions: The number of alternative courses of action, or choices, for a decision.

Output Options:

Title: The custom title for the worksheet generated by this procedure.

Notes:

The Optimum Actions cells will initially display the #N/A message. These messages will disappear when you completely fill in the payoff table.

7   Probability & Prob. Distributions

7.1   Simple & Joint Probabilities

Creates a new worksheet that calculates probabilities from a 2 x 2 cross-classification, or contingency, table of outcomes. The worksheet generated contains an empty table of outcomes into which you enter data for each event to complete the analysis. You can also enter labels for each event in the sample space.

7.2 Normal Probability Distribution

Creates a new worksheet that calculates normal probabilities.

Data:

Mean: The population mean.

Standard Deviation: The population standard deviation.

Input Options:

Probability for: X <=: If selected, worksheet calculates the probability of values less than or equal to the entered X value.

Probability for: X >: If selected, worksheet calculates the probability of values greater than the entered X value.

Probability for range: If selected, worksheet calculates the probability of values between the entered range of X values.

X for Cumulative Percentage: If selected, worksheet calculates the X and Z values corresponding to a cumulative area under the normal curve.

Output Options:

Title: The custom title for the worksheet generated by this procedure.

Notes:

If the Probability for: X<= and Probability for: X > are both selected, the worksheet also calculates the sum of these probabilities.

7.3 Normal Probability Plot

Creates a normal probability plot on a new chart sheet.

Data:

Variable Cell Range: The single-column cell range containing the numerical data to be summarized.

First cell contains label: If selected, the contents of the first cell of the variable cell range is treated as a descriptive label and not as a data value.

Output Options:

Title: The custom title to be placed on the chart sheet generated.

Notes:

This procedure also generates a worksheet that calculates the Z values for each value plotted.

7.4 Binomial Probability Distribution

Creates a table of binomial probabilities on a new worksheet.

Data:

Sample Size: The sample size.

Prob. of an Event of Interest: The probability of an event of interest ("success").

Outcomes From:/To: The range of the number of successes in the sample to include in the probability table. An outcome range from 0 to the sample size generates the entire table.

Output Options:

Title: The custom title for the worksheet generated by this procedure.

Cumulative Probabilities: If selected, adds additional table columns that calculate cumulative probabilities for P(<=X), P(<X), P(>X), and P(>=X).

Histogram: If selected, generates a histogram, based on the table of binomial probabilities, on a separate chart sheet.

Notes:

The outcomes range should include non-negative values that do not exceed the sample size. If the range includes values greater than the sample size, #NUM! error values will displayed for those range values greater than the sample size.

Given the limitations of the Microsoft Excel function used to calculate binomial probabilities, the values in the table should be considered as fair approximations of the actual probability values.

7.5   Exponential Probability Distribution

Calculates the exponential probability for values less than or equal to a given X value on a new worksheet.

Data:

Mean per unit (Lambda): The mean of the exponential distribution.

X Value: The X value of interest.

Output Options:

Title: The custom title for the worksheet generated by this procedure.

7.6   Hypergeometric Probability Distribution

Creates a table of hypergeometric probabilities on a new worksheet.

Data:

Sample Size: The sample size.

No. of Events of Interest in Pop.: The number of events of interest ("successes") in the population.

Population Size: The population size.

Output Options:

Title: The custom title for the worksheet generated by this procedure.

Histogram: If selected, generates a histogram, based on the table of hypergeometric probabilities, on a separate chart sheet.

Notes:

If the value for the number of successes in the population is larger than the sample size (an impossibility), #NUM! error values will displayed for at least some of the cells of the table and you should ignore all values in the table until you enter a proper value for the number of successes in the population.

7.7 Poisson Probability Distribution

Creates a table of Poisson probabilities for 0 to 20 successes per unit (X) on a new worksheet.

Data:

Mean/Expected No. of Events of Interest: The mean or expected number of events per unit (*lambda*).

Output Options:

Title: The custom title for the worksheet generated by this procedure.

Cumulative Probabilities: If selected, adds additional table columns that calculate cumulative probabilities for P(<=X), P(<X), P(>X), and P(>=X).

Histogram: If selected, generates a histogram, based on the table of Poisson probabilities, on a separate chart sheet.

Notes:

Given the limitations of the Microsoft Excel function used to calculate Poisson probabilities, the values in the table should be considered as fair approximations of the actual probabilities values.

8 Sampling Procedures

8.1 Sampling Distributions Simulation

Creates a simulated sampling distribution of the mean on a new worksheet.

Data:

Number of Samples: The number of samples to generate.

Sample Size: The sample size.

Distribution Options:

Uniform: If selected, samples are generated from a uniformly distributed population between 0 and 1.

Standardized Normal: If selected, samples are generated from standardized normally distributed population.

Discrete: If selected, samples are generated from a discrete distribution that is defined by the X and P(X) values cell range.

X and P(X) Values Cell Range: The two-column cell range containing the X and P(X) values that define the discrete distribution. This cell range should *not* contain labels in the first row. (Enabled only if Discrete is selected.)

Output Options:

Title: The custom title for the worksheet generated by this procedure.

Histogram: If selected, generates a histogram, based on the simulated distribution, on a separate chart sheet.

9   Confidence Intervals Procedures

9.1   Estimate for the Mean, sigma known

Calculates the confidence interval estimate for the mean when sigma is known on a new worksheet.

Data:

Population Standard Deviation: The population standard deviation.

Confidence Level: The confidence level as a percentage. Default value is 95%.

Input Options:

Sample Statistics Known: If selected, procedure uses the sample statistics entered in the dialog box.

    Sample Size: The sample size. (Enabled only if Sample Statistics Known is selected.)

    Sample Mean: The sample mean. (Enabled only if Sample Statistics Known is selected.)

Sample Statistics Unknown: If selected, sample statistics are calculated from the data values in the sample cell range.

    Sample Cell Range: The single-column cell range containing the numerical data of the sample. (Enabled only if Sample Statistics Unknown is selected.)

    First cell contains label: If selected, the contents of the first cell of the sample cell range is treated as a descriptive label and not as a data value. (Enabled only if Sample Statistics Unknown is selected.)

Output Options:

Title: The custom title for the worksheet generated by this procedure.

Finite Correction Factor: If selected, adds rows to the new worksheet that calculate the finite population correction factor.

    Population Size: The population size. (Enabled only if Finite Correction Factor is selected.)

9.2   Estimate for the Mean, sigma unknown

Calculates the confidence interval estimate for the mean when sigma is unknown on a new worksheet.

Data:

Confidence Level: The confidence level as a percentage. Default value is 95%.

Input Options:

Sample Statistics Known: If selected, procedure uses the sample statistics entered in the dialog box.

    Sample Size: The sample size. (Enabled only if Sample Statistics Known is selected.)

    Sample Mean: The sample mean. (Enabled only if Sample Statistics Known is selected.)

    Sample Std. Deviation: The sample standard deviation. (Enabled only if Sample Statistics Known is selected.)

Sample Statistics Unknown: If selected, sample statistics are calculated from the data values in the sample cell range.

    Sample Cell Range: The single-column cell range containing the numerical data of the sample. (Enabled only if Sample Statistics Unknown is selected.)

First cell contains label: If selected, the contents of the first cell of the sample cell range is treated as a descriptive label and not as a data value. (Enabled only if Sample Statistics Unknown is selected.)

Output Options:

Title: The custom title for the worksheet generated by this procedure.

Finite Correction Factor: If selected, adds rows to the new worksheet that calculate the finite population correction factor.

Population Size: The population size. (Enabled only if Finite Correction Factor is selected.)

9.3 Estimate for the Population Variance

Calculates the confidence interval estimate for the population variance and standard deviation on a new worksheet.

Data:

Sample Size: The sample size.

Sample Standard Deviation: The sample standard deviation.

Confidence Level: The confidence level as a percentage. Default value is 95%.

Output Options:

Title: The custom title for the worksheet generated by this procedure.

9.4 Estimate for the Proportion

Calculates the confidence interval estimate for the population proportion on a new worksheet.

Data:

Sample Size: The sample size.

Number of Successes: The number of successes.

Confidence Level: The confidence level as a percentage. Default value is 95%.

Output Options:

Title: The custom title for the worksheet generated by this procedure.

Notes:

If the value of the sample size is less than the number of successes (an impossibility), #NUM! error messages will appear in the cells containing the confidence interval lower and upper limits.

9.5 Estimate for the Population Total

Calculates the confidence interval estimate for the population total on a new worksheet.

Data:

Population Size: The population size.

Confidence Level: The confidence level as a percentage. Default value is 95%.

Input Options:

Sample Statistics Known: If selected, procedure uses the sample statistics entered in the dialog box.

Sample Size: The sample size. (Enabled only if Sample Statistics Known is selected.)

Sample Mean: The sample mean. (Enabled only if Sample Statistics Known is selected.)

Sample Standard Deviation: The sample standard deviation. (Enabled only if Sample Statistics Known is selected.)

Sample Statistics Unknown: If selected, sample statistics are calculated from the data values in the sample cell range.

Sample Cell Range: The single-column cell range containing the numerical data of the sample. (Enabled only if Sample Statistics Unknown is selected.)

First cell contains label: If selected, the contents of the first cell of the sample cell range is treated as a descriptive label and not as a data value. (Enabled only if Sample Statistics Unknown is selected.)

Output Options:

Title: The custom title for the worksheet generated by this procedure.

9.6   Estimate for the Total Difference

Calculates the confidence estimate for the total difference on a new worksheet.

Data:

Sample Size: The sample size.

Population Size: The population size.

Confidence Level: The confidence level as a percentage. Default value is 95%.

Differences Cell Range: The single-column cell range containing the non-zero differences.

First cell contains label: If selected, the contents of the first cell of the sample cell range is treated as a descriptive label and not as a data value.

Output Options:

Title: The custom title for the worksheet generated by this procedure.

Notes:

This procedure also generates a second worksheet that contains the data necessary to calculate the sum of differences and sum of squares for the non-zero differences.

10   Sample Size Procedures

10.1   Sample Size Determination for the Mean

Calculates the sample size needed for estimating the population mean on a new worksheet.

Data:

Population Standard Deviation: The population standard deviation.

Sampling Error: The acceptable sampling error.

Confidence Level: The confidence level as a percentage. Default value is 95%.

Output Options:

Title: The custom title for the worksheet generated by this procedure.

Finite Correction Factor: If selected, adds rows to the new worksheet that calculate the finite

population correction factor.

Population Size: The population size. (Enabled only if Finite Correction Factor is selected.)

10.2 Sample Size Determination for the Proportion

Calculates the sample size needed for estimating the population proportion on a new worksheet.

Data:

Estimate of True Proportion: The estimate of the true proportion of success.

Sampling Error: The acceptable sample error.

Confidence Level: The confidence level as a percentage. Default value is 95%.

Output Options:

Title: The custom title for the worksheet generated by this procedure.

Finite Correction Factor: If selected, adds rows to the new worksheet that calculate the finite population correction factor.

Population Size: The population size. (Enabled only if Finite Correction Factor is selected.)

Notes:

This procedure also generates a second worksheet that contains the data necessary to calculate the sum of differences and sum of squares for the non-zero differences.

11 One-Sample Procedures

11.1 Z Test for the Mean, sigma known

Performs the Z test of hypothesis for the mean when sigma is known on a new worksheet.

Data:

Null Hypothesis: The null hypothesis value.

Level of Significance: The level of significance.

Population Standard Deviation: The population standard deviation.

Input Options:

Sample Statistics Known: If selected, procedure uses the sample statistics entered in the dialog box.

Sample Size: The sample size. (Enabled only if Sample Statistics Known is selected.)

Sample Mean: The sample mean. (Enabled only if Sample Statistics Known is selected.)

Sample Statistics Unknown: If selected, sample statistics are calculated from the data values in the sample cell range.

Sample Cell Range: The single-column cell range containing the numerical data of the sample. (Enabled only if Sample Statistics Unknown is selected.)

First cell contains label: If selected, the contents of the first cell of the sample cell range is treated as a descriptive label and not as a data value. (Enabled only if Sample Statistics Unknown is selected.)

Test Options:

Two-Tail Test: If selected, worksheet performs a two-tail test.

Upper-Tail Test: If selected, worksheet performs an upper-tail test.

Lower-Tail Test: If selected, worksheet performs a lower-tail test.

Output Options:

Title: The custom title for the worksheet generated by this procedure.

11.2   t Test for the Mean, sigma unknown

Performs the t test of hypothesis for the mean when sigma is unknown on a new worksheet.

Data:

Null Hypothesis: The null hypothesis value.

Level of Significance: The level of significance.

Input Options:

Sample Statistics Known: If selected, procedure uses the sample statistics entered in the dialog box.

Sample Size: The sample size. (Enabled only if Sample Statistics Known is selected.)

Sample Mean: The sample mean. (Enabled only if Sample Statistics Known is selected.)

Sample Standard Deviation: The sample standard deviation. (Enabled only if Sample Statistics Known is selected.)

Sample Statistics Unknown: If selected, sample statistics are calculated from the data values in the sample cell range.

Sample Cell Range: The single-column cell range containing the numerical data of the sample. (Enabled only if Sample Statistics Unknown is selected.)

First cell contains label: If selected, the contents of the first cell of the sample cell range is treated as a descriptive label and not as a data value. (Enabled only if Sample Statistics Unknown is selected.)

Test Options:

Two-Tail Test: If selected, worksheet performs a two-tail test.

Upper-Tail Test: If selected, worksheet performs an upper-tail test.

Lower-Tail Test: If selected, worksheet performs a lower-tail test.

Output Options:

Title: The custom title for the worksheet generated by this procedure.

11.3   Chi-Square Test for the Variance

Performs the Chi-Square test for the variance on a new worksheet.

Data:

Null Hypothesis: The null hypothesis value.

Level of Significance: The level of significance.

Sample Size: The sample size

Sample Standard Deviation: The sample standard deviation

Test Options:

Two-Tail Test: If selected, worksheet performs a two-tail test.

Upper-Tail Test: If selected, worksheet performs an upper-tail test.

Lower-Tail Test: If selected, worksheet performs a lower-tail test.

Output Options:

Title: The custom title for the worksheet generated by this procedure.

11.4   Z Test for the Proportion

Performs the *Z* test of hypothesis for the proportion on a new worksheet.

Data:

Null Hypothesis: The null hypothesis value.

Level of Significance: The level of significance.

Number of Items of Interest: The number of items of interest in the sample. Also sometimes known as the number of successes.

Sample Size: The sample size.

Test Options:

Two-Tail Test: If selected, worksheet performs a two-tail test.

Upper-Tail Test: If selected, worksheet performs an upper-tail test.

Lower-Tail Test: If selected, worksheet performs a lower-tail test.

Output Options:

Title: The custom title for the worksheet generated by this procedure.

12   Two-Sample Procedures (Unsummarized Data)

12.1   Pooled-Variance t Test (unsummarized data)

Performs the pooled-variance *t* test of hypothesis for differences between two means on a new worksheet.

Data:

Hypothesized Difference: The hypothesized difference between the means of the two populations.

Level of Significance: The level of significance. Default value is 0.05.

Population 1 Sample Cell Range: The single-column cell range containing the sample data from Population 1.

Population 2 Sample Cell Range: The single-column cell range containing the sample data from Population 2.

First cells in both ranges contain label: If selected, the contents of the first cell of each sample cell range is treated as a descriptive label and not as a data value.

Test Options:

Two-Tail Test: If selected, worksheet performs a two-tail test.

Upper-Tail Test: If selected, worksheet performs an upper-tail test.

Lower-Tail Test: If selected, worksheet performs a lower-tail test.

Output Options:

Title: The custom title for the worksheet generated by this procedure.

Confidence Interval Estimate: If selected, procedure calculates the confidence interval estimate of the difference between two means (of two independent populations).

   Confidence level: The confidence level as a percentage. (Enabled only if Confidence Interval Estimate is selected.)

12.2   Separate-Variance t Test (unsummarized data)

Performs the separate-variance *t* test of hypothesis for differences between two means on a new worksheet.

Data:

Hypothesized Difference: The hypothesized difference between the means of the two populations.

Level of Significance: The level of significance. Default value is 0.05.

Population 1 Sample Cell Range: The single-column cell range containing the sample data from Population 1.

Population 2 Sample Cell Range: The single-column cell range containing the sample data from Population 2.

First cells in both ranges contain label: If selected, the contents of the first cell of each sample cell range is treated as a descriptive label and not as a data value.

Test Options:

Two-Tail Test: If selected, worksheet performs a two-tail test.

Upper-Tail Test: If selected, worksheet performs an upper-tail test.

Lower-Tail Test: If selected, worksheet performs a lower-tail test.

Output Options:

Title: The custom title for the worksheet generated by this procedure.

12.3   Paired t Test (unsummarized data)

Performs the paired *t* test of hypothesis for differences between two means on a new worksheet.

Data:

Hypothesized Mean Difference: The hypothesized difference between the means of the two populations.

Level of Significance: The level of significance. Default value is 0.05.

Population 1 Sample Cell Range: The single-column cell range containing the sample data from Population 1.

Population 2 Sample Cell Range: The single-column cell range containing the sample data from Population 2.

First cells in both ranges contain label: If selected, the contents of the first cell of each sample cell range is treated as a descriptive label and not as a data value.

Test Options:

Two-Tail Test: If selected, worksheet performs a two-tail test.

Upper-Tail Test: If selected, worksheet performs an upper-tail test.

Lower-Tail Test: If selected, worksheet performs a lower-tail test.

Output Options:

Title: The custom title for the worksheet generated by this procedure.

12.4   F Test for Differences in Two Variances (unsummarized data)

Performs the *F* test of hypothesis for differences between two variances on a new worksheet.

Data:

Level of Significance: The level of significance. Default value is 0.05.

Population 1 Sample Cell Range: The single-column cell range containing the sample data from Population 1.

Population 2 Sample Cell Range: The single-column cell range containing the sample data from Population 2.

First cells in both ranges contain label: If selected, the contents of the first cell of each sample cell range is treated as a descriptive label and not as a data value.

Test Options:

Two-Tail Test: If selected, worksheet performs a two-tail test.

Upper-Tail Test: If selected, worksheet performs an upper-tail test.

Output Options:

Title: The custom title for the worksheet generated by this procedure.

Notes:

On the worksheet created, the procedure rearranges the two samples as "Larger-Variance Sample" and "Smaller-Variance Sample." Because of the asymmetric shape of the *F* distribution, this eliminates the need for a Lower-Tail Test option.

12.5   Wilcoxon Rank Sum Test

Performs the Wilcoxon rank sum test for differences between two medians on a new worksheet.

Data:

Level of Significance: The level of significance. Default value is 0.05.

Population 1 Sample Cell Range: The single-column cell range containing the sample data from Population 1.

Population 2 Sample Cell Range: The single-column cell range containing the sample data from Population 2.

First cells in both ranges contain label: If selected, the contents of the first cell of each sample cell range is treated as a descriptive label and not as a data value.

Test Options:

Two-Tail Test: If selected, worksheet performs a two-tail test.

Upper-Tail Test: If selected, worksheet performs an upper-tail test.

Lower-Tail Test: If selected, worksheet performs a lower-tail test.

Output Options:

Title: The custom title for the worksheet generated by this procedure.

Notes:

This procedure also generates a second worksheet that contains the ranks of the sample data from the two populations.

13   Two-Sample Procedures (Summarized Data)

13.1   Pooled-Variance t Test

Performs the pooled-variance *t* test of hypothesis for differences between two means on a new worksheet.

Data:

Hypothesized Difference: The hypothesized difference between the means of the two populations.

Level of Significance: The level of significance. Default value is 0.05.

Population 1 Sample:

Sample Size: The sample size for the Population 1 sample.

Sample Mean: The sample mean for the Population 1 sample.

Sample Standard Deviation: The sample standard deviation for the Population 1 sample.

Population 2 Sample:

Sample Size: The sample size for the Population 2 sample.

Sample Mean: The sample mean for the Population 2 sample.

Sample Standard Deviation: The sample standard deviation for the Population 2 sample.

Test Options:

Two-Tail Test: If selected, worksheet performs a two-tail test.

Upper-Tail Test: If selected, worksheet performs an upper-tail test.

Lower-Tail Test: If selected, worksheet performs a lower-tail test.

Output Options:

Title: The custom title for the worksheet generated by this procedure.

Confidence Interval Estimate: If selected, procedure calculates the confidence interval estimate of the difference between two means (of two independent populations).

Confidence level: The confidence level as a percentage. (Enabled only if Confidence Interval Estimate is selected.)

13.2   Separate-Variance t Test

Performs the separate-variance *t* test of hypothesis for differences between two means on a new worksheet.

Data:

Hypothesized Difference: The hypothesized difference between the means of the two populations.

Level of Significance: The level of significance. Default value is 0.05.

Population 1 Sample:

Sample Size: The sample size for the Population 1 sample.

Sample Mean: The sample mean for the Population 1 sample.

Sample Standard Deviation: The sample standard deviation for the Population 1 sample.

Population 2 Sample:

Sample Size: The sample size for the Population 2 sample.

Sample Mean: The sample mean for the Population 2 sample.

Sample Standard Deviation: The sample standard deviation for the Population 2 sample.

Test Options:

Two-Tail Test: If selected, worksheet performs a two-tail test.

Upper-Tail Test: If selected, worksheet performs an upper-tail test.

Lower-Tail Test: If selected, worksheet performs a lower-tail test.

Output Options:

Title: The custom title for the worksheet generated by this procedure.

13.3 Paired t Test

Performs the paired *t* test of hypothesis for differences between two means on a new worksheet.

Data:

Hypothesized Mean Difference: The hypothesized difference between the means of the two populations.

Level of Significance: The level of significance. Default value is 0.05.

Differences Cell Range: The single-column cell range containing the differences.

First cell contains label: If selected, the contents of the first cell of the differences cell range is treated as a descriptive label and not as a data value.

Test Options:

Two-Tail Test: If selected, worksheet performs a two-tail test.

Upper-Tail Test: If selected, worksheet performs an upper-tail test.

Lower-Tail Test: If selected, worksheet performs a lower-tail test.

Output Options:

Title: The custom title for the worksheet generated by this procedure.

13.4 Z Test for Differences in Two Means

Performs the *Z* test of hypothesis for differences between two means on a new worksheet.

Data:

Hypothesized Difference: The hypothesized difference between the means of the two populations.

Level of Significance: The level of significance. Default value is 0.05.

Population 1 Sample:

Sample Size: The sample size for the Population 1 sample.

Sample Mean: The sample mean for the Population 1 sample.

Population Std. Deviation: The population standard deviation for the Population 1 sample.

Population 2 Sample:

    Sample Size: The sample size for the Population 2 sample.

    Sample Mean: The sample mean for the Population 2 sample.

    Population Std. Deviation: The population standard deviation for the Population 2 sample.

Test Options:

    Two-Tail Test: If selected, worksheet performs a two-tail test.

    Upper-Tail Test: If selected, worksheet performs an upper-tail test.

    Lower-Tail Test: If selected, worksheet performs a lower-tail test.

Output Options:

    Title: The custom title for the worksheet generated by this procedure.

13.5   F Test for Differences in Two Variances

Performs the *F* test of hypothesis for differences between two variances on a new worksheet.

Data:

    Level of Significance: The level of significance. Default value is 0.05.

Population 1 Sample:

    Sample Size: The sample size for the Population 1 sample.

    Sample Standard Deviation: The sample standard deviation for the Population 1 sample.

Population 2 Sample:

    Sample Size: The sample size for the Population 2 sample.

    Sample Standard Deviation: The sample standard deviation for the Population 2 sample.

Test Options:

    Two-Tail Test: If selected, worksheet performs a two-tail test.

    Upper-Tail Test: If selected, worksheet performs an upper-tail test.

Output Options:

    Title: The custom title for the worksheet generated by this procedure.

Notes:

On the worksheet created, the procedure rearranges the two samples as "Larger-Variance Sample" and "Smaller-Variance Sample." Because of the asymmetric shape of the *F* distribution, this eliminates the need for a Lower-Tail Test option.

13.6   Chi-Square Test for Differences in Two Proportions

Creates a new worksheet that performs the Chi-square test of hypothesis for differences between two proportions. The worksheet generated contains an empty observed frequencies table into which you enter the observed frequencies to complete the analysis. (You can also enter labels for the row and column variables.)

Data:

    Level of Significance: The level of significance. Default value is 0.05.

Output Options:

    Title: The custom title for the worksheet generated by this procedure.

Notes:

Some worksheet cells will initially display the #DIV/0! message. These messages will disappear when you completely fill in the observed frequency table.

13.7 Z Test for the Difference in Two Proportions

Performs the *Z* test of hypothesis for differences between two proportions on a new worksheet.

Data:

Hypothesized Difference: The hypothesized difference between the proportions of the two populations.

Level of Significance: The level of significance. Default value is 0.05.

Population 1 Sample:

Number of Items of Interest: The number of items of interest in the Population 1sample. Also sometimes known as the number of successes.

Sample Size: The sample size for the Population 1 sample.

Population 2 Sample:

Number of Items of Interest: The number of items of interest in the Population 2 sample. Also sometimes known as the number of successes.

Sample Size: The sample size for the Population 2 sample.

Test Options:

Two-Tail Test: If selected, worksheet performs a two-tail test.

Upper-Tail Test: If selected, worksheet performs an upper-tail test.

Lower-Tail Test: If selected, worksheet performs a lower-tail test.

Output Options:

Title: The custom title for the worksheet generated by this procedure.

Confidence Interval Estimate: If selected, procedure calculates the confidence interval estimate of the difference between two proportions.

Confidence level: The confidence level as a percentage. (Enabled only if Confidence Interval Estimate is selected.)

13.8 McNemar Test

Creates a new worksheet that performs the McNemar test of hypothesis for differences between two proportions (related samples). The worksheet generated contains an empty observed frequencies table into which you enter the observed frequencies to complete the analysis. (You can also enter labels for the row and column variables.)

Data:

Level of Significance: The level of significance. Default value is 0.05.

Test Options:

Two-Tail Test: If selected, worksheet performs a two-tail test.

Upper-Tail Test: If selected, worksheet performs an upper-tail test.

Lower-Tail Test: If selected, worksheet performs a lower-tail test.

Output Options:

Title: The custom title for the worksheet generated by this procedure.

Notes:

Some worksheet cells will initially display the #DIV/0! Message.

14 Multiple-Sample Procedures

14.1 Chi-Square Test

Creates a new worksheet that performs the Chi-square test of hypothesis for differences in the proportions among multiple independent populations. The worksheet generated contains an empty cross-classification, or contingency, table into which you the observed frequencies to complete the analysis. (You can also enter labels for the row and column variables.)

Data:

Level of Significance: The level of significance.

Number of Rows: The number of rows.

Number of Columns: The number of columns, i.e., independent populations.

Output Options:

Title: The custom title for the worksheet generated by this procedure.

Marascuilo Procedure: Performs the Marascuilo procedure for comparisons between all pairs of groups (enabled if 2 is entered as the Number of Rows).

Notes:

Many worksheet cells will initially display the #DIV/0! message. These messages will disappear when you completely fill in the observed frequency table.

14.2 Kruskal-Wallis Rank Test

Performs the Kruskal-Wallis rank sum test of hypothesis for differences between medians from multiple independent sample groups on a new worksheet.

Data:

Level of Significance: The level of significance.

Sample Data Cell Range: The multiple-column cell range containing the sample data, arranged one column per independent sample. Range must contain at least three contiguous columns.

First cells contain label: If selected, the contents of the first cell of each column in the sample data cell range is treated as a descriptive label and not as a data value.

Output Options:

Title: The custom title for the worksheet generated by this procedure.

Notes:

This procedure also generates a second worksheet that contains the ranks of the sample data from the multiple independent samples.

Any non-numeric values in the sample data cell range (excluding the first cells of columns if First

cells contain label is selected) are treated as zero values for the purposes of ranking.

14.3 Levene Test

Performs the Levene Test for homogeneity of variance for multiple independent sample groups on a new worksheet

Data:

Level of Significance: The level of significance.

Sample Data Cell Range: The multiple-column cell range containing the sample data, arranged one column per independent sample. Range must contain at least two contiguous columns.

First cells contain label: If selected, the contents of the first cell of each column in the sample data cell range is treated as a descriptive label and not as a data value.

Output Options:

Title: The custom title for the worksheet generated by this procedure.

14.4 One-Way ANOVA

Performs an one-way ANOVA on a new worksheet.

Data:

Level of Significance: The level of significance.

Group Data Cell Range: The multiple-column cell range containing the sample data, arranged one column per independent group. Range must contain at least three contiguous columns.

First cells contain label: If selected, the contents of the first cell of each column in the sample data cell range is treated as a descriptive label and not as a data value.

Output Options:

Title: The custom title for the worksheet generated by this procedure.

Tukey-Kramer Procedure: Inserts a new worksheet that performs the Tukey-Kramer procedure (to determine which means from multiple independent groups are significantly different). You will need to enter the *Q* statistic into this worksheet to complete the analysis

Notes

This procedure also generates a second worksheet that contains ANOVA statistics that are referenced in the Tukey-Kramer worksheet.

14.5 Two-Way ANOVA with replication

Performs a two-factor ANOVA with replications on a new worksheet.

Data:

Level of Significance: The level of significance. Default value is 0.05.

Population 1 Sample Cell Range: The single-column cell range containing the sample data from Population 1.

Population 2 Sample Cell Range: The single-column cell range containing the sample data from Population 2.

First cells in both ranges contain label: If selected, the contents of the first cell of each

sample cell range is treated as a descriptive label and not as a data value.

Output Options:

Title: The custom title for the worksheet generated by this procedure.

Cell Means Plot: Creates a cell means plot on a separate chart sheet.

15   Control Charts Procedures

15.1   p Chart

Creates a *p* Chart on a new chart sheet.

Data:

Nonconformances Cell Range: The single-column cell range containing the nonconformance numerical data.

First cell contains label: If selected, the contents of the first cell in the nonconformances cell range is treated as a descriptive label and not as a data value.

Sample/Subgroups:

Size does not vary: If selected, the sample/subgroup size for all subgroups taken is the same and is equal to the value entered in the dialog box for the sample/subgroup size.

Sample/Subgroup Size: The sample/subgroup size for all subgroups. (Enabled only if Size does not vary is selected.)

Size varies: If selected, the sample/subgroup size for each subgroup is determined by the values in the sample/subgroup size cell range.

Sample/Subgroup Size Cell Range: The single-column cell range containing the sample/subgroup sizes for each subgroup taken. (Enabled only if Size varies is selected.)

First cell contains label: If selected, the contents of the first cell of the sample/subgroup size cell range is treated as a descriptive label and not as a data value. (Enabled only if Size varies is selected.)

Output Options:

Title: The custom title for the chart sheet generated by this procedure.

Notes:

This procedure also generates two worksheets that contain calculations for the control limits and data for the *p* chart.

The procedure will display a warning message if at least one subgroup sample size differs too much from the mean sample size.

15.2   c Chart

Creates a *c* Chart on a new chart sheet.

Data:

Nonconformances Cell Range: The single-column cell range containing the nonconformance numerical data.

First cell contains label: If selected, the contents of the first cell in the nonconformances cell range is treated as a descriptive label and not as a data value.

Output Options:

Title: The custom title for the chart sheet generated by this procedure.

Notes:

This procedure also generates two worksheets that contain calculations for the control limits and data for the *c* chart.

15.3   R and XBar Charts

Creates an *R* chart on a new chart sheet. Optionally generates an *XBar* chart on another chart sheet.

Data:

Subgroup/Sample Size: The subgroup/sample size.

Subgroup Ranges Cell Range: The single-column cell range containing the subgroup range data for each subgroup taken.

First cell contains label: If selected, the contents of the first cell in the subgroup range cell range is treated as a descriptive label and not as a data value.

Chart Options:

R Chart Only: If selected, the procedure generates an R chart.

R and XBar Charts: If selected, the procedure generates an R and an XBar chart and uses the subgroup mean data for each subgroup found in the subgroup means cell range.

   Subgroup Means Cell Range: The single-column cell range containing the subgroup means for each subgroup taken. (Enabled only if R and XBar Charts is selected.)

   First cell contains label: If selected, the contents of the first cell of the sample/subgroup size cell range is treated as a descriptive label and not as a data value. (Enabled only if R and XBar Charts is selected.)

Output Options:

Title: The custom title for the chart sheet generated by this procedure.

Notes:

This procedure also generates two worksheets that calculations for the control limits and data for the chart(s).

16   Regression Procedures

16.1   Simple Linear Regression

Creates a simple linear regression analysis on a new worksheet.

Data:

Y Variable Cell Range: The single-column cell range containing the values of the Y variable.

X Variable Cell Range: The single-column cell range containing the values of the X variable. Range must be on the same worksheet as the Y variable cell range.

First cells in both ranges contain label: If selected, the contents of the first cell in each cell range treated as a descriptive label and not as a data value

Confidence level for the regression coefficients: The confidence level as a percentage. Default value is 95%.

Regression Tool Output Options:

Regression Statistics Table: If selected, worksheet includes a table of regression statistics. (Selected by default)

ANOVA and Coefficients Table: If selected, worksheet includes ANOVA and coefficients table.

Residuals Table: If selected, worksheet includes a table of residual values.

Residual Plot: If selected, procedure generates a residual plot on a new chart sheet.

Output Options:

Title: The custom title for the worksheet generated by this procedure.

Scatter Plot: If selected, procedure generates a scatter plot of the Y and X values.

Durbin-Watson Statistic: If selected, procedure calculates the Durbin-Watson statistic on a separate worksheet.

Confidence Int. Est. and Prediction Int. for X=: If selected, procedure calculates the confidence interval estimate for the average predicted Y and prediction interval for the individual response Y on a separate worksheet using the X value entered next to this selection and confidence level value entered in the dialog box.

Confidence level for interval: The confidence level as a percentage. (Enabled only if **Confidence Int. Est. and Prediction Int. for X= i**s selected.)

Notes:

This procedure will create a number of supporting worksheets and chart sheets, depending on the options selected.

Some output options will force the selection of certain regression tool output options.

Data that are perfectly correlated will cause #NUM! and $DIV/0! error messages in the regression results worksheet.

16.2   Multiple Regression

Creates a multiple regression analysis on a new worksheet.

Data:

Y Variable Cell Range: The single-column cell range containing the values of the Y variable.

X Variables Cell Range: The multiple-column cell range containing the values of the X variable. This range must be on the same worksheet as the Y variable cell range. Unlike most PHStat2 cell ranges, this range can include more than one area (non-adjacent columns) so long as all columns are on the same worksheet.

First cells in both ranges contain label: If selected, the contents of the first cell of every column in each cell range is treated as a descriptive label and not as a data value

Confidence level for the regression coefficients: The confidence level as a percentage. Default value is 95%.

Regression Tool Output Options:

Regression Statistics Table: If selected, worksheet includes a table of regression statistics. (Selected by default)

ANOVA and Coefficients Table: If selected, worksheet includes ANOVA and coefficients table.

> **Residuals Table:** If selected, worksheet includes a table of residual values.

> **Residual Plots:** If selected, procedure generates residual plots on new chart sheets.

Output Options:

> **Title:** The custom title for the worksheet generated by this procedure.

> **Durbin-Watson Statistic:** If selected, procedure calculates the Durbin-Watson statistic on a separate worksheet.

> **Variance Inflationary Factor (VIF):** If selected, procedure calculates the VIF for each regression of a single *X* variable and all other *X* variables on separate worksheets.

> **Confidence Interval Estimates & Prediction Interval:** If selected, procedure calculates the confidence interval estimates for the mean predicted Y and the prediction interval for the individual response Y on a separate worksheet using the confidence level value entered in the dialog box.

>> **Confidence level for intervals:** The confidence level as a percentage. (Enabled only if **Confidence Interval Estimates & Prediction Interval** is selected.)

Notes:

This procedure will create a number of supporting worksheets and chart sheets, depending on the options selected.

Some output options will force the selection of certain regression tool output options.

Data for an *X* variable that are perfectly correlated will cause all other *X* variable coefficients to be reported as 0.

16.3   Best Subsets

Creates a best subsets multiple regression analysis on a new worksheet.

Data:

> **Y Variable Cell Range:** The single-column cell range containing the values of the Y variable.

> **X Variables Cell Range:** The multiple-column cell range containing the values of the X variable. This range must be on the same worksheet as the Y variable cell range.

> **First cells in both ranges contain label:** If selected, the contents of the first cell of every column in each cell range is treated as a descriptive label and not as a data value

> **Confidence level for regression coefficients:** The confidence level as a percentage. Default value is 95%.

Output Options:

> **Title:** The custom title for the worksheet generated by this procedure.

Notes:

This procedure generates additional sheets, each containing the results of one of the regression analyses considered.

This procedure is limited to 7 *X* (independent) variables.

16.4   Stepwise Regression

Creates a stepwise regression analysis on a new worksheet.

Data:

> **Y Variable Cell Range:** The single-column cell range containing the values of the Y variable.
>
> **X Variables Cell Range:** The multiple-column cell range containing the values of the X variable. This range must be on the same worksheet as the Y variable cell range. Unlike most PHStat2 cell ranges, this range can include more than one area (non-adjacent columns) so long as all columns are on the same worksheet.
>
> **First cells in both ranges contain label:** If selected, the contents of the first cell of every column in each cell range is treated as a descriptive label and not as a data value
>
> **Confidence level for regression coefficients:** The confidence level as a percentage. Default value is 95%.

Stepwise Criteria:

> **p values:** If selected, p values are used as the stepwise criteria.
>
> **t values:** If selected, t values are used as the stepwise criteria.

Stepwise Options:

> **General Stepwise:** If selected, the procedure uses the general stepwise method to enter and remove variables using the p or t values entered in the dialog box as its criteria.
>
> > **p/t value to enter** and **p/t value to remove:** criteria values for the general stepwise method. (Enabled only if General Stepwise is selected.)
>
> **Forward Selection:** If selected, the procedure uses the forward selection method to enter variables using the p or t value entered in the dialog box.
>
> > **p/t value to enter:** criteria value for the forward selection method. (Enabled only if Forward Selection is selected.)
>
> **Backward Elimination:** If selected, the procedure uses the backward elimination method to remove variables using the p or t value entered in the dialog box.
>
> > **p/t value to remove:** criteria value for the backward elimination method. (Enabled only if Backward Elimination is selected.)

Output Options:

> **Title:** The custom title for the worksheet generated by this procedure.

Notes:

This procedure also generates on a separate worksheet a multiple regression analysis using all of the X variables.

This procedure handles up to 7 X Variables. If your X Variables Cell Range that contains more than 7 columns (7 X variables), this procedure will display a message asking you to rerun the procedure with 7 or fewer X variables.

17   Utility Procedures

17.1   Fix Up Chart

Modifies chart, eliminating some of the formatting errors that Excel makes when it creates a chart.

17.2 Remove Worksheet Cell Tints

Removes the light turquoise and light yellow tints of data and results cells in worksheets generated by PHStat.

Index

- A -

Addition rule   17

- B -

Backward Elimination   39
Bar Chart   14
Best Subsets   38
Binomial Probability Distribution   18
Box-and-Whisker Plot   10
Boxplot   10

- C -

c Chart   35
Chi-Square Test   33
Chi-Square Test for Differences in Two Proportions
31
Chi-Square Test for the Variance   25
Covariance   16
Covariance and Portfolio Management   16
Cumulative Percentage Polygon   12

- D -

Data preparation   7
Dot Scale Diagram   11
Durbin-Watson Statistic   36, 37

- E -

Estimate for the Mean sigma known   21
Estimate for the Mean sigma unknown   21
Estimate for the Population Total   22
Estimate for the Population Variance   22
Estimate for the Proportion   22
Estimate for the Total Difference   23
Expected Monetary Value   16
Expected Opportunity Loss   17
Expected placeOpportunity Loss   17
Exponential Probability Distribution   19

- F -

F Test for Differences in Two Variances   31
F Test for Differences in Two Variances unsummarized
data   28
Five-Number Summary   10
Fix Up Chart   39
For More Information   9
Forward Selection   39
Frequency Distribution   12
Frequency Polygon   12

- G -

General Stepwise   39
Getting Started with PHStat2   8

- H -

Histogram   12
Histogram & Polygons   12
How PHStat2 presents its results   7
Hypergeometric   19
Hypergeometric Probability Distribution   19

- K -

Kruskal-Wallis Rank Test   33

- L -

Levene s Test   34

- M -

Marascuilo procedure   33
McNemar Test   32
Microsoft Excel   6
Microsoft Excel PivotTable   14, 15
Multiple Regression   37

- N -

New Statistical Procedures   5