NOTE E: GUIDE TO EXCEL STATISTICAL FUNCTIONS, ROUTINES AND TOOLS

1. GENERAL INFORMATION

(1) Excel 2000, 2003 and 2007 support 128 statistical and related functions, which are listed below. These functions and routines have been in all three versions, with some minor differences between each of the three versions.

(2) Logic on interpretation of inputs

(a) Normally functions will skip cells with text (string) and skip cells with TRUE or FALSE (logical) and will also skip blank cells (empty).

(b) Functions that end in "A" interpret cells with text as the number 0, cells with a TRUE as the number 1, cells with FALSE as the number 0, and cells with numbers as the number entered. Empty or blank cells are skipped..

(c) Functions that end in "P" divide the sum by the count of cells (n) by n rather than n-1 (which is normal).

(d) Functions ending in "PA" combine both (b) and (c) operations on the cells.

(3) Certain functions beginning with D, (DAVERAGE, DCOUNT, DCOUNTA, DMAX, DMIN, DSTDEV, DSTDEVP, DVAR and DVARP) do the same statistical operations corresponding to the remaining letters, but are database oriented and require 3 inputs, "database", "field" and "criteria". Database specifies the range of cells describing the data base, field indicates which column is used in the function, and criteria is a range of cells that indicate the selection criteria. The examples under HELP are the best way to understand how these functions work, and how criteria is linked.

(4) A function always has at least one pair of parentheses, even if there is no input.

2. STATISTICAL AND RELATED FUNCTIONS AND ROUTINES COMMONLY USED

Name of Function/Routine	Source	Type	Class
ANOVA: Single Factor	Analysis Tool-Pac	Routine	ANOVA
ANOVA: Two-Factor With Replication	Analysis Tool-Pac	Routine	ANOVA
ANOVA: Two-Factor Without Replication	Analysis Tool-Pac	Routine	ANOVA
Correlation	Analysis Tool-Pac	Routine	Correlations
Covariance	Analysis Tool-Pac	Routine	Correlations
Descriptive Statistics	Analysis Tool-Pac	Routine	Univariate
Exponential Smoothing	Analysis Tool-Pac	Routine	
F-test: Two Sample for Variances	Analysis Tool-Pac	Routine	Test of Significance
Fourier Analysis	Analysis Tool-Pac	Routine	Math
Histogram	Analysis Tool-Pac	Routine	Univariate

Moving Average	Analysis Tool-Pac	Routine	
Random Number Generator	Analysis Tool-Pac	Routine	RNG
Rank and Percentage	Analysis Tool-Pac	Routine	Univariate
Regression	Analysis Tool-Pac	Routine	Regression
Sampling	Analysis Tool-Pac	Routine	RNG
t-Test: Paired Two Sample for Means	Analysis Tool-Pac	Routine	Test of Significance
t-Test: Two-sample Assuming Equal Variance	Analysis Tool-Pac	Routine	Test of Significance
t-Test: Two-sample Assuming Unequal Variances	Analysis Tool-Pac	Routine	Test of Significance
z-Test: Two Sample for Means	Analysis Tool-Pac	Routine	Test of Significance
AVEDEV	Excel Basic	Function	Univariate
AVERAGE	Excel Basic	Function	Univariate
AVERAGEA	Excel Basic	Function	Univariate
COUNT	Excel Basic	Function	Univariate
COUNTA	Excel Basic	Function	Univariate
COUNTBLANK	Excel Basic	Function	Univariate
COUNTIF	Excel Basic	Function	Univariate
DEVSQ	Excel Basic	Function	Univariate
DSTDEV	Excel Basic	Function	Univariate
DSTDEVP	Excel Basic	Function	Univariate
DVAR	Excel Basic	Function	Univariate
DVARP	Excel Basic	Function	Univariate
GEOMEAN	Excel Basic	Function	Univariate
HARMEAN	Excel Basic	Function	Univariate
KURT	Excel Basic	Function	Univariate
LARGE	Excel Basic	Function	Univariate
MAX	Excel Basic	Function	Univariate
MAXA	Excel Basic	Function	Univariate
MEDIAN	Excel Basic	Function	Univariate
MIN	Excel Basic	Function	Univariate
MINA	Excel Basic	Function	Univariate
MODE	Excel Basic	Function	Univariate
PERCENTILE	Excel Basic	Function	Univariate
PERCENTRANK	Excel Basic	Function	Univariate
QUARTILE	Excel Basic	Function	Univariate
RANK	Excel Basic	Function	Univariate
SKEW	Excel Basic	Function	Univariate
SMALL	Excel Basic	Function	Univariate
STDEV	Excel Basic	Function	Univariate
STDEVA	Excel Basic	Function	Univariate
STDEVP	Excel Basic	Function	Univariate
STDEVPA	Excel Basic	Function	Univariate
TRIMEAN	Excel Basic	Function	Univariate
VAR	Excel Basic	Function	Univariate
VARA	Excel Basic	Function	Univariate
VARP	Excel Basic	Function	Univariate
VARPA	Excel Basic	Function	Univariate
FREQUENCY	Excel Basic	Array	Univariate

BETADIST	Excel Basic	Function	Distribution
BETAINV	Excel Basic	Function	Distribution
BINOMDIST	Excel Basic	Function	Distribution
CRITBINOM	Excel Basic	Function	Distribution
CHIDIST	Excel Basic	Function	Distribution
CHINV	Excel Basic	Function	Distribution
EXPONDIST	Excel Basic	Function	Distribution
FDIST	Excel Basic	Function	Distribution
FINV	Excel Basic	Function	Distribution
GAMMADIST	Excel Basic	Function	Distribution
GAMMAINV	Excel Basic	Function	Distribution
HYPGEOMDIST	Excel Basic	Function	Distribution
LOGNORMDIST	Excel Basic	Function	Distribution
LOGINV	Excel Basic	Function	Distribution
NEGBINOMDIST	Excel Basic	Function	Distribution
NORDIST	Excel Basic	Function	Distribution
NORMINV	Excel Basic	Function	Distribution
NORMSDIST	Excel Basic	Function	Distribution
NORMSINV	Excel Basic	Function	Distribution
POISSON	Excel Basic	Function	Distribution
TDIST	Excel Basic	Function	Distribution
TINV	Excel Basic	Function	Distribution
WEIBULL	Excel Basic	Function	Distribution
CHITEST	Excel Basic	Function	Test of Significance
FTEST	Excel Basic	Function	Test of Significance
TTEST	Excel Basic	Function	Test of Significance
ZTEST	Excel Basic	Function	Test of Significance
CORREL	Excel Basic	Function	Correlations
COVAR	Excel Basic	Function	Correlations
PEARSON	Excel Basic	Function	Correlations
RSQ	Excel Basic	Function	Correlations
FORECAST	Excel Basic	Function	Regression
INTERCEPT	Excel Basic	Function	Regression
SLOPE	Excel Basic	Function	Regression
STEYX	Excel Basic	Function	Regression
LINEST	Excel Basic	Array	Regression
TREND	Excel Basic	Array	Regression
GROWTH	Excel Basic	Array	Regression
LOGEST	Excel Basic	Array	Regression
CONFIDENCE	Excel Basic	Function	Misc
COMBIN	Excel Basic	Function	Math
FISHER	Excel Basic	Function	Misc
FISHERINV	Excel Basic	Function	Misc
GAMMALN	Excel Basic	Function	Math
PERMUT	Excel Basic	Function	Math
RAND	Excel Basic	Function	RNG
RANDBETWEEN	Excel Basic	Function	RNG

STANDARDIZE	Excel Basic	Function	Misc
PROB	Excel Basic	Function	Misc
EXP	Excel Basic	Function	Math
FACT	Excel Basic	Function	Math
FACTDOUBLE	Excel Basic	Function	Math
LN	Excel Basic	Function	Math
LOG10	Excel Basic	Function	Math
COLUMNS	Excel Basic	Function	Spreadsheet
MULTINOMIAL	Excel Basic	Function	Math
ROWS	Excel Basic	Function	Spreadsheet
SUM	Excel Basic	Function	Math
SUMPRODUCT	Excel Basic	Function	Matrix
SUMSQ	Excel Basic	Function	Math
SUMX2MY2	Excel Basic	Function	Math
SUMX2PY2	Excel Basic	Function	Math
SUMSMY2	Excel Basic	Function	Math
SUMIF	Excel Basic	Function	Math
SUMIFS	Excel Basic	Function	Math
COUNTIF	Excel Basic	Function	Math
MDETERM	Excel Basic	Function	Matrix
MINVERSE	Excel Basic	Array	Matrix
MMULT	Excel Basic	Array	Matrix
TRANSPOSE	Excel Basic	Array	Matrix
SOLVER	Add-In	Routine	Misc

3. EXPANDED INFORMATION ON THE FUNCTIONS AND ROUTINES

A. DATA CALCULATIONS

Output values recalculated each time the worksheet is recalculated.

UNIVARIATE FUNCTIONS, RANGE INPUT, SINGLE CELL OUTPUT

AVEDEV – Returns the average of the absolute deviations of data points from their mean.

AVERAGE – Returns the average of a list of data points (or arguments), numbers only.

AVERAGEA – AVERAGE with the “A” input.

COUNT – Counts how many numbers there are in a range.

COUNTA – COUNT with the “A” input. The COUNTA function skips only blank cells

COUNTBLANK – Counts the number of blank (empty) cells in a range.

COUNTIF – The COUNT function where a cell in the range input has to meet a single criteria (number, logical expression or text) in order to be counted.

DEVSQ – Returns the sum of squares of deviations from the mean of a range of data values.

DSTDEV – Database Function (database, field, criteria). Returns the standard deviation estimate of a population based on a sample range of numbers. Divides by n-1.
DSTDEVP – DSTDEV with the "P" computation.
DVAR – Database Function (database, field, criteria). Returns the variance estimate of a population based on a sample range of numbers. Divides by n-1.
DVARP – DVAR with the "P" computation.
GEOMEAN – Returns the geometric mean of a range of data points.
HARMEAN – Returns the harmonic mean of a range of data points
KURT – Returns the kurtosis value of a data set
LARGE – Returns the k-th largest value in a data set
MAX – Returns the maximum value in a list of data points (or arguments), numbers only.
MAXA – MAX with the "A" input.
MEDIAN – Returns the median of a list of data points
MIN - Returns the minimum value in a list of data points (or arguments), numbers only
MINA – MIN with the "A" input.
MODE – Returns the most common (frequent) value in a data set
PERCENTILE – Returns the k-th percentile of values in a range of data points
PERCENTRANK – Returns the percentage rank of a value in a data set
QUARTILE – Returns a selected quartile value (min, first, median, third, or max) from a data set
RANK – Returns the rank of a selected data point (number) from a list of data points
SKEW – Returns the skewness value of a data set.
SMALL – Returns the k-th smallest value from a data set.
STDEV – Returns the standard deviation estimate of a population based on a sample data set of numbers. Divides by n-1.
STDEVA –STDEV with the "A" input.
STDEVP– STDEV with the "P" computation.
STDEVPA – STDEV with the "PA" option.
TRIMEAN – Returns the trimmed mean of a selected interior range of a data set.
VAR – Returns the variance estimate of a population based on a sample data set of numbers.
VARA – VAR with the "A" input.
VARP– Returns the variance of a population based on a set of numbers defining the population.
VARPA – VAR with the "PA" option.

UNIVARIATE FUNCTIONS, RANGE OUTPUT

{FREQUENCY} – Returns a frequency distribution as a vertical array of bins.

B. DISTRIBUTIONS

Output values recalculated each time the worksheet is recalculated.

PARAMETER INPUTS, SINGLE CELL OUTPUT

DISTRIBUTION	DENSITY	CUMMULATIVE	INVERSE
Beta		BETADIST	BETAINV
Binomial	BINOMDIST	CRITBINOM	
Chi-Square		CHIDIST	CHINV
Exponential	EXPONDIST	EXPONDIST	
F		FDIST	FINV
Gamma	GAMMADIST	GAMMADIST	GAMMAINV
Hyper geometric	HYPGEOMDIST		
Log Normal		LOGNORMDIST	LOGINV
Negative Binomial	NEGBINOMDIST		
Normal (with parameters)	NORMDIST	NORMDIST	NORMINV
Normal (z values)		NORMSDIST	NORMSINV
Poisson	POISSON		
T		TDIST	TINV
Weibull		WEIBULL	

C. TESTS

Output values recalculated each time the worksheet is recalculated.

RANGE AND PARAMETER INPUTS, SINGLE CELL OUTPUT

CHITEST (Function) – This is a Chi-square Goodness-of-Fit test for grouped data. It does not support general Chi Square tests on variances. The test will only work on 2 way contingency tables. The test cannot be applied to single lists of observed and expected values. The first input, "actual range" is the range of the observed values, as a 2-way contingency table. The second input is "expected range", the range of a separate contingency table giving the expected values.

FTEST (Function) - Returns the one-tailed probability value of an F test on two separate ranges of data. The ranges may be of different lengths.

TTEST (Function) - Returns the probability value of a t test on two separate data sets. Function allows for 1 or 2 tail tests, paired data and equal-unequal variances

ZTEST (Function) - Returns the two-tailed probability of a normal distribution z test on a range of data with respect to a known population mean and standard deviation. If the standard deviation field is left blank, the routine uses the standard deviation of the data.

D. CORRELATIONS

Output values recalculated each time the worksheet is recalculated.

RANGE INPUTS, SINGLE CELL OUTPUT

CORREL - Returns the correlation coefficient between two separate ranges of paired data.

COVAR - Returns the covariance between two separate ranges of paired data.

PEARSON – Returns the Pearson product moment correlation between two separate ranges of paired data

RSQ - Returns the squared Pearson product moment correlation between two separate ranges of paired data

E. CURVE FITTING - (LINEAR REGRESSION)

RANGE AND PARAMETER INPUTS, SINGLE CELL OUTPUT

FORECAST – Given an X range of data and a Y range of data, returns a predicted Y value for a given X value.

INTERCEPT – Returns the intercept of a least mean squares straight line through an imputed range of X and Y values.

SLOPE – Returns the slope of a least mean squares straight line through an imputed range of X and Y values

STEYX - Returns the standard error of the residuals from a least mean squares straight line through an imputed range of X and Y values

RANGE INPUTS, RANGE OUTPUT

Calculated only when the function is called.

{LINEST} – Fits a least mean squares straight line to a vector of Y values and a range of X values. Returns a block of coefficient values, an intercept value, and regression statistics in a certain order. Multivariate (up to 16 x variables) can be fitted, and by transformations, polynomial and some non-linear models can be fitted.

{TREND} – Fits a least mean squares line to a given range of X and Y values, and returns a range of predicted new Y values for an input range of new X values based on the fitted line. Accepts multivariate X ranges.

F. CURVE FITTING - (NON-LINEAR REGRESSION)

Calculated only when the function is called.

{GROWTH} – Fits an exponential curve to a given range of X and Y values, and returns a range of predicted new Y values for an input range of new X values based on the fitted curve. Accepts multivariate X ranges. Fits linear regression to the logs of the X values and transforms back

{LOGEST} – Fits an exponential function to a vector of Y values and a range of X values and returns a vector of exponential values, a coefficient, and regression statistics in a certain order. It is the application of LINEST on the logs of Y and X values, and conversion to an exponential system.

G. OTHER

Output values recalculated each time the worksheet is recalculated.

PARAMETER INPUTS, SINGLE CELL OUTPUT

CONFIDENCE – Returns the confidence interval about a sample mean that will include the population mean when the population true standard deviation is known.

COMBIN – Returns the number of ways that n2 objects can be selected from n1 objects, without regard for order. Integer inputs and outputs.

FISHER – Returns the Fisher transformation of a correlation coefficient (close to 1) that approximates normal distribution z values.

FISHERINV – Returns a correlation coefficient, which is the inverse of the Fisher transformation, given a z value.

GAMMALN – Returns the natural logarithm of the gamma function

PERMUT – Returns the number of permutations of a set of chosen objects from a larger set of objects.

RAND – Returns a uniformly distributed random number between 0 and 1.

RANDBETWEEN – Returns a uniformly distributed random number between a lower and upper input parameter.

STANDARDIZE - Returns a calculated standardized z value of a data point from a data set that has a mean and a standard deviation.

RANGE AND PARAMETER INPUTS, SINGLE CELL OUTPUT

PROB – Returns the probability that values in a range of X are between a lower and upper limit, given probability values for each of the X values and bounds on the value for which the probability value corresponds to.

2. USEFUL SUPPORTING FUNCTIONS

Output values recalculated each time the worksheet is recalculated.

A. SINGLE PARAMETER INPUT, SINGLE VALUE OUTPUT

EXP – Returns e to the power of a number.

FACT – Returns the factorial of a number (integer only).

FACTDOUBLE – Returns the double factorial of a number (integer only).

LN – Returns the natural logarithm of a number

LOG10 – Returns the base 10 logarithm of a number

B. RANGE INPUT, SINGLE VALUE OUTPUT

COLUMNS – Returns the number of columns in a range

MULTINOMIAL – Returns the ratio of the factorial of a sum of numbers (numerator) in a range to the product of the factorial for each number (denominator) in the range

ROWS – Returns the number of rows in a range

SUM – Returns the sum of numbers in a range.

SUMPRODUCT – Returns a single value, which is the sum of the product of two ranges. Useful to obtain the dot, scalar or inner product of two vectors. The outer product is a matrix, which can be obtained from MMULT.

SUMSQ – Returns the sum of squares of the numbers in a range.

SUMX2MY2 – Returns the sum of the difference between the squares of paired values in two separate ranges.

SUMX2PY2 – Returns the sum of the sum of the squares of paired values in two separate ranges

SUMXMY2 – Returns the sum of the square of the difference between paired numbers from two separate ranges.

SUMIF- First tests each cell in the range, and only adds it in if the criteria is met. (one criteria)

SUMIFS - First tests each cell in the range, and only adds it in if the criteria is met. (multiple criteria)

COUNTIF - First tests each cell in the range, and only counts it if the criteria is met.

C. LOGIC FUNCTIONS

IF – The IF function returns one of two values depending on the evaluation of criteria as to being true or false. There are three arguments, which can be nested if functions. Argument 1, a logical expression (test) that results in either a logical TRUE or FALSE. Argument 2, the value returned (which can be an equation) if the test results in TRUE. Argument 3, the value returned (which can be an equation) if the test results in FALSE.

AND – Returns a TRUE if all arguments are TRUE, else returns FALSE.

OR – Returns a TRUE if at least one argument is TRUE, else returns FALSE.

NOT – Returns the complement of its argument.

D. SPECIAL FUNCTIONS

PI() – Returns 3.14159….. No input parameters

E. MATRIX OPERATIONS, RANGE INPUTS, SINGLE OR RANGE OUTPUTS

MDETERM – Returns the matrix determinant (as a single value) of a symmetric array. All cells within the range must have a number.

{MINVERSE} – Returns the inverse matrix of a symmetric matrix. All cells within the range must have a number.

{MMULT} – Returns the matrix product of two arrays. The two arrays must have conformal rows and columns. Both arrays must have numbers in all cells.

{TRANSPOSE} – Returns the transpose of an array. Used to shift col-row arrays to row-col form

IMPORTANT: Names enclosed in {} require the ctrl-shift-enter operation for array entry

3. TOOLS - SOLVER

Assumes that the Solver add-in has been loaded

A Solver ready worksheet must first be prepared before starting Solver.

Solver Parameters (menu): Specifies cells with parameters and the cell with the final equation.

Set Target Cell: _____A cell with a formula that calculates the final value. For example this may be a cell that sums the square of the residuals in a column. Another column would carry the equation for calculating a Y value for each point. Etc.

Equal to: Select Max, Min or a Value to be reached by the calculated value in the target cell.

By Changing Cells: _____A range that includes the parameter cells that SOLVER can change. If this is left blank and **Guess** is selected, Solver will attempt to change all cells on the sheet that have no formulas, but have values. It is critical that these variable parameter cells contain values reasonably close to an acceptable solution. They should not be left empty.

Subject to the Constraints: Push Add, Change, or Delete to enter and modify constraints.

Add Constraint:

Cell Reference: Enter the cell location of the parameter to be constrained
Logic: Select the desired logic reference to a constraint value
Constraint: Enter a value
OK Button: Return to the Solver Main Window
Cancel: Deletes the constraint shown
Add: Select to add another constraint
Help:
Change Constraint: Shows the selected constraint where changes can be made.
Cell Reference:
Logic: Select the desired logic reference to a constraint value
Constraint: Change a value
OK Button: Return to the Solver Main Window
Cancel: Deletes the constraint shown
Add: Select to add another constraint
Help:

Solve: Select this button to start solver.

Solver Results: Solver returns information about the solution. These could be, “Solver found a solution. All constraints and optimality conditions are satisfied”, ”Solver could not find a feasible solution”, ”The maximum iteration limit was reached, continue anyway?”, “The maximum time limit was reached, continue anyway?”
Keep Solver Solution or Restore Original Values: Select one.
Reports: Select one if desired.
Answer Report: Lists the target cell and displays both the original and changed values of the adjustable cells. Information about the constraints that were applied against those cells is also available.
Sensitivity Report: Provides information about the sensitivity level of the target cell to changes in your constraints.
Limits Report: Displays the values, upper and lower limits, and target values of the target and adjustable cells. The limits are the values

(lowest and highest) the cell can accept and still satisfy the terms of your constraints.

OK: Return to the Solver Main Window

Cancel:

Save Scenario: Select to save results.

Scenario Name: Enter a title/name for the solution

OK: Return to Solver Results:

Cancel: Cancels this selection

Help:

Help:

Close: Select this button to exit solver

Options: Select this button to get the Solver Options dialog box

Options (menu): See **Help** for details

Max Time: _____ Seconds

Iterations:_____

Precision:_____

Tolerance:_____%

Convergence:_____

Assume Linear Model:____

Assume Non-Negative:____

Use Automatic Scaling:____

Show Iteration Results:_____

Estimates, ___Target or ____Quadratic

Derivatives, ___Forward or _____Central

Search, ____Newton or _____Conjugate

OK:

Cancel:

Load Model:

Save Model:

Help:

Reset All: Select this button to restart Solver from the beginning..

Help:

4. TOOLS – DATA ANALYSIS ROUTINES

Assumes that the Data Analysis ToolPak add-in has been loaded

A. GENERAL:

Each of these routines calculates values only once, when they are invoked. If input values change, you have to rerun the routine as a brand-new routine from the menu.

Data inputs are defined as ranges. Select the red icon at the right side of the input box. The menu disappears and the input line stands alone. Use the mouse and the left button to select the input block (data range), and the correct input appears in the stand-alone input box. If correct, select the red icon again, and the previous menu appears.

All The routines have the following standard output selection, however some will only allow the Output Range option.

Standard Output Options: (Select one from the following 3 options)

Output Range: (Input the coordinates of the starting cell for the output on the current active worksheet)
New Worksheet Ply: (Input a new worksheet name. Name Appears on the tab of the new worksheet.)
New Workbook: (Establish file name when the file is saved)

B. ROUTINES:

1. Anova: Single Factor: Performs single-factor analysis of variance
Input Range: A block defining the input data
Grouped By Columns or Rows: Do each group of replicates occur as a row or as a column. Select one.
Labels in First Row: Does the first cell in each group have a label?.
Alpha Value: The p value of the tail for critical F test values.

2. Anova: Two-Factor with Replication: Performs two-factor analysis of variance with replication.
Input Range: A block defining the input data. The input data has to be specially placed in the input block. The first row (top) is reserved for column labels (column effects). The first column (right) is reserved for labels (row effects). The replicates appear as a contiguous group of rows under one row effect. The input data block must include the top label row and the side label column.
Rows per Sample: The number of replicates per group (row effects) as a number of rows.
Alpha Value: The p value of the tail for critical F test values.

3. Anova: Two Factor without Replication: Performs two-factor analysis of variance without replication.
Input Range: A block defining the input data. Each column is effect 1, each row is effect 2.
Labels in First Row: If checked, range has labels in the top (first) row and labels in the first (left) column.
Alpha Value: The p value of the tail for critical F test values.

4. Correlation:. Computes correlation coefficients between variables
Input Range: A block defining the input data. Two or more variables.
Grouped By Columns or Rows: Are the individual X values of a variable along a row or down a column. Select one.
Labels in First Row: Does the first cell in each variable group have a label?
Output: An NxN block with the lower triangle filled with correlation coefficients and with diagonal values.= 1.

5. Covariance: Computes covariances between variables

Input Range: A block defining the input data. Two or more variables.
Grouped By Columns or Rows: Are the individual X values of a variable along a row or down a column. Select one.
Labels in First Row: Does the first cell in each variable group have a label?.
Output: An NxN block with the lower triangle filled with covariance values and with diagonal values equal to the variance of each variable.

6. Descriptive Statistics: A table of univariate statistic values
Input Range: A block defining the input data. A list of values.
Grouped By Columns or Rows: Is the list along a row or down a column. Select one.
Labels in First Row: Does the first cell in the list have a label?
Summary Statistics: Check box for summary statistics
Confidence Level for Mean: Check box for +/- confidence level about the mean and put the desired level in the next box. **(Percentage):**
Kth Largest: Check box for the kth largest value and enter the **K value:**
Kth Smallest: Check box for the kth smallest value and enter the **K value:**
Output:
Mean:
Standard Error:
Median:
Mode:
Standard Deviation:
Sample Variance:
Kurtosis:
Skewness:
Range:
Minimum:
Maximum:
Sum:
Count:
Largest(k):
Smallest(k):
Confidence Level(Pct):

7. Exponential Smoothing: Smooths a ragged series of numbers. Useful on time varying data.
Input Range: A block defining the input data. A list of values.
Dampening Factor: A number
Labels: Does the first cell in the list have a label?
Chart Output: Check box for a chart
Standard Errors: Check box for the chart to show error bars.

8. F-Test Two-Sample for Variances:
Input Variable 1 Range: A block defining the input data. A list of values.
Input Variable 2 Range: A block defining the input data. A list of values.

Labels: Does the first cell in the list have a label?
Alpha?

9. Fourier Analysis: Use when you have knowledge of the mathematics of Fourier Transforms

Input Range: A block defining the input data. A list of values. (not exceeding 4096 values)
Labels in First Row: Does the first cell in the list have a label?
Inverse: Check box for an inverse analysis output.

10. Histogram:

Input Range: A block defining the input data. A list of values.
Input Bin Range: A block defining the bin values. A list of equally spaced values.
Labels: Does the first cell in the input data list have a label?
Pareto (sorted histogram): Check box for this option
Cumulative Percentage: Check box for this option
Chart Output: Check box for this option

11. Moving Average: Smooths a time varying series.

Input Range: A block defining the input data. A list of values.
Labels in First Row: Does the first cell in the list have a label?
Interval: A number
Chart Output: Check box for a chart
Standard Errors: Check box for the chart to show error bars.

12. Random Number Generator.[1]

Menu's indicate the appropriate inputs for each of the following distributions. The output goes to a defined range of columns and rows, defined as a number of variables (columns) and a number of values (rows). Total number of generations is M = columns * rows, except for the patterned case. Represents a sample of M values from a defined population. The output is referred to as a block below.

[1] The random number generator used to generate values is different from RAND. The generator used here is defective and has a short period. It is very granular. Has a very limited range. It comes from a 16 bit integer. It should not be used for any serious work.

Uniform

Produces a block of M random numbers (real floating point) between a lower limit and an upper limit, where the numbers in-between are uniformly distributed within this interval.

Normal (*Will not produce z values less than 5 or more than 5.)*

Produces a block of M random numbers (real floating point) from a normal distribution with a given population mean and standard deviation.

Bernoulli

Produces a block of M binary numbers (integer), either a 1 or a 0, with the frequency of the 1's defined by the input p value. Occurrence of the 1 or 0 is random.

Binomial

Produces a block of M random numbers (integer) of "good" items in N trials when p is the probability of the "good" item occurring for one event.

Poisson

Produces a block of M random numbers (integer) of events that will occur based on a lambda value, where lambda is the population mean value or when lambda is the product of a rate constant C times a time interval T, and where the number of events are for the time interval T.

Patterned.

Produces a block of increasing non-random repeating numbers (real floating point) from A to B in steps of C that looks like a sawtooth. Each number can be repeated n consecutive times in a sequence. Produces only repeating sequences of increasing numbers. The starting number in a sequence is A and the last number in a sequence is the value B The input of the number of sequences defines the number in each row, and it overrules the set in number of values. The block size is variable to obtain an integer number of sequences in each column.

Discrete

Produces a block of M random discrete numbers (either integer or floating point) defined from a two column fixed table array. This table has in the first column a list of the discrete values, and in the second column a probability value for the occurrence of each discrete value. The probability values in the second column must add up to one. There is no interpolation between the discrete numbers. Represents a random output from a histogram.

13. Rank and Percentile: Not recommended. Output confusing.

Input Range: A block defining the input data. A list of values.

Grouped By Columns or Rows: Is the list along a row or down a column. Select one.

Labels in First Row: Does the first cell in the list have a label?
Output:
Column 1, Point: The rank of the point in the original data list
Column 2: Label name for the data, sorted in descending order
Column 3: Rank number for that point in the sorted list. Given tied data, the rank number is the rank in the full set when the new value first occurs.
Column 4: Percentage (unknown what this means)

14. Regression:

Input Y Range: A block defining the input Y data. A list of values.
Input X Range: A block defining the input X data. For a single variable it is a list, for multivariate data it is a contiguous block of variables.
Labels: Does the first cell in the input data have a label? Applied to both X and Y data blocks.
Constant is Zero: Check box if the regression is forced to go through the origin.
Confidence Level: Check box for +/- confidence level about the computed coefficient values and put the desired level in the next box. **(Percentage).**
Residuals: Boxes to be checked are:
Residuals: Check box if a column of residuals is wanted.
Standardized Residuals: Check if a column of standardized residuals is wanted.
Residual Plots: Check box if plot is desired
Line Fit Plots: Check box if plot is desired
Normal Probability Plots: Check box if plot is desired.
Output: Extensive, occupies 9 columns by 20 or more rows. Overwrites any cells. Include a row for each variable, 6 rows for any residual table headers and a row of residual values for each data point

15. Sampling:

Input Range: A block defining the input data. A list of values.
Labels: Does the first cell in the list have a label?
Periodic or Random Sampling: Select one.
Periodic: The input range is sampled every nth cell. The value of n goes in the box.
Random[2]: The input range is randomly sampled giving a sample of n cells. The number of random samples goes in the box.

TWO SAMPLE TESTS (These explanations are directly from HELP)

The Two-Sample t-Test analysis tools test for equality of the population means underlying each sample. The three tools employ different assumptions: that the population variances are equal, that the population variances are not equal, and

[2] The random number generator used to generate values is different from RAND. The generator used here is defective and has a short period. It is very granular. Has a very limited range. It comes from a 16 bit integer. It should not be used for any serious work.

that the two samples represent before treatment and after treatment observations on the same subjects.

For all three tools below, a t-Statistic value, t, is computed and shown as "t Stat" in the output tables. Depending on the data, this value, t, can be negative or non-negative. Under the assumption of equal underlying population means, if t < 0, "P(T <= t) one-tail" gives the probability that a value of the t-Statistic would be observed that is more negative than t. If t >=0, "P(T <= t) one-tail" gives the probability that a value of the t-Statistic would be observed that is more positive than t. "t Critical one-tail" gives the cutoff value so that the probability of observing a value of the t-Statistic greater than or equal to "t Critical one-tail" is Alpha.

"P(T <= t) two-tail" gives the probability that a value of the t-Statistic would be observed that is larger in absolute value than t. "P Critical two-tail" gives the cutoff value so that the probability of an observed t-Statistic larger in absolute value than "P Critical two-tail" is Alpha.

16. t-Test: Paired Two Sample for Means:

Input Variable 1 Range: A block defining the input data. A list of values.
Input Variable 2 Range: A block defining the input data. A list of values.
Hypothesized Mean Difference: A number
Labels: Does the first cell in both lists have a label?
Alpha: The probability of rejection of the hypothesis.

Output:	Variable 1	Variable 2
Mean	X	X
Variance	X	X
Observations	X	X
Pearson Correlation	X	
Hypothesized Mean Difference	X	
Df	X	
T Stat	X (Computed t value)	
P(T<=t) one-tail	X	
t Critical one-tail	X (For the given alpha value)	
P(T<=t) two-tail	X (The test is T not equal to t)	
t Critical two-tail	X (For the given alpha value)	

t-Test: Paired Two Sample For Means You can use a paired test when there is a natural pairing of observations in the samples, such as when a sample group is tested twice— before and after an experiment. This analysis tool and its formula perform a paired two-sample student's t-test to determine whether observations taken before a treatment and observations taken after a treatment are likely to have come from distributions with equal population means. This t-test form does not assume that the variances of both populations are equal.

Note: Among the results generated by this tool is pooled variance, an accumulated measure of the spread of data about the mean, derived from the following formula.

17. t-Test: Two Sample Assuming Equal Variances:

Input Variable 1 Range: A block defining the input data. A list of values.
Input Variable 2 Range: A block defining the input data. A list of values.
Hypothesized Mean Difference: A number
Labels: Does the first cell in both lists have a label?
Alpha: The probability of rejection of the hypothesis.

Output:	Variable 1	Variable 2
Mean	X	X
Variance	X	X
Observations	X	X
Pooled Variance	X	
Hypothesized Mean Difference	X	
Df	X	
T Stat	X (Computed t value)	
P(T<=t) one-tail	X	
t Critical one-tail	X (For the given alpha value)	
P(T<=t) two-tail	X (The test is T not equal to t)	
t Critical two-tail	X (For the given alpha value)	

t-Test: Two-Sample Assuming Equal Variances This analysis tool performs a two-sample student's t-test. This t-test form assumes that the two data sets came from distributions with the same variances. It is referred to as a homoscedastic t-test. You can use this t-test to determine whether the two samples are likely to have come from distributions with equal population means.

18. t-Test: Two-Sample Assuming Unequal Variances:

Input Variable 1 Range: A block defining the input data. A list of values.
Input Variable 2 Range: A block defining the input data. A list of values.
Hypothesized Mean Difference: A number
Labels: Does the first cell in both lists have a label?
Alpha: The probability of rejection of the hypothesis.

Output:	Variable 1	Variable 2
Mean	X	X
Variance	X	X
Observations	X	X
Hypothesized Mean Difference	X	
Df	X	
T Stat	X (Computed t value)	
P(T<=t) one-tail	X	
t Critical one-tail	X (For the given alpha value)	
P(T<=t) two-tail	X (The test is T not equal to t)	
t Critical two-tail	X (For the given alpha value)	

t-Test: Two-Sample Assuming Unequal Variances This analysis tool performs a two-sample student's t-test. This t-test form assumes that the two data sets came

from distributions with unequal variances. It is referred to as a heteroscedastic t-test. As with the Equal Variances case above, you can use this t-test to determine whether the two samples are likely to have come from distributions with equal population means. Use this test when the there are distinct subjects in the two samples. Use the Paired test, described below, when there is a single set of subjects and the two samples represent measurements for each subject before and after a treatment.

19. z-Test: Two Sample for Means:

Input Variable 1 Range: A block defining the input data. A list of values.
Input Variable 2 Range: A block defining the input data. A list of values.
Hypothesized Mean Difference: A number
Variable 1 Variance (known):
Variable 2 Variance (known):
Labels: Does the first cell in both lists have a label?
Alpha: The probability of rejection of the hypothesis.

Output:	Variable 1	Variable 2
Mean	X	X
Variance	X	X
Observations	X	X
Hypothesized Mean Difference	X	
z	X (Computed z value)	
P(Z<=z) one-tail	X	
z Critical one-tail	X (For the given alpha value)	
P(Z<=z) two-tail	X (The test is Z not equal to z)	
z Critical two-tail	X (For the given alpha value)	

The z-Test: Two Sample for Means analysis tool performs a two-sample z-test for means with known variances. This tool is used to test the null hypothesis that there is no difference between two population means against either one-sided or two-sided alternative hypotheses. If variances are not known, the worksheet function, ZTEST, should be used instead.

When using the z-Test tool, one should be careful to understand the output. “P(Z <= z) one-tail” is really P(Z >= ABS(z)), the probability of a z-value further from 0 in the same direction as the observed z value when there is no difference between the population means. “P(Z <= z) two-tail” is really P(Z >= ABS(z) or Z <= -ABS(z)), the probability of a z-value further from 0 in either direction than the observed z-value when there is no difference between the population means. The two-tailed result is just the one-tailed result multiplied by 2. The z-Test tool can also be used for the case where the null hypothesis is that there is a specific non-zero value for the difference between the two population means.

5. TRENDLINES

These are overlay lines put on the chart that represent a fit of an equation to the plotted data. These include exponential, linear, logarithmic, polynomial, or power equations. Also just a moving average can be shown as a trendline. The methods to put in a trendline are different for Excel 2007 when compared to the Excel 2003 method.

The fitted equation with parameter values can be shown on the chart.

The algorithms used here for calculating trendline fits are different from LINEST and LOGEST. Fitted parameter values will be slightly different.