

# Data Analysis - Random Sampling

<http://cameron.econ.ucdavis.edu/excel/ex01access.html>  
<http://www.betersolutions.com/default.aspx>

## ADDING-IN THE DATA ANALYSIS TOOLPACK TO EXCEL

Statistical analysis such as descriptive statistics and regression requires the Excel Data Analysis add-in. The default configuration of Excel **does not** automatically support descriptive statistics and regression analysis. **You may need to add these** to your computer (a once-only operation).

**Excel 2007:** The Data Analysis add-in should appear at right-end of Data menu as Data Analysis. If not then

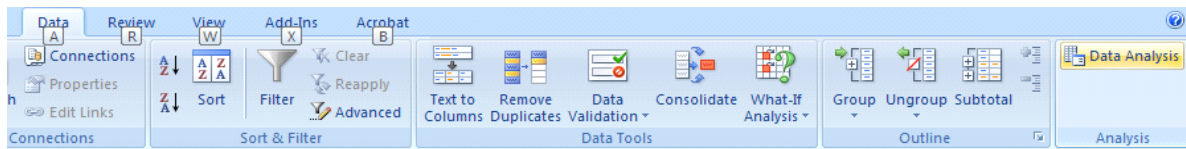
1. Click the **Microsoft Office Button**, and then click **Excel Options**.
2. Click **Add-Ins**, and then in the **Manage** box, select **Excel Add-ins**.
3. Click **Go**.
4. In the **Add-Ins available** box, select the **Analysis ToolPak** check box, and then click **OK**.  
**Tip** If **Analysis ToolPak** is not listed in the **Add-Ins available** box, click **Browse** to locate it.  
If you get prompted that the Analysis ToolPak is not currently installed on your computer, click **Yes** to install it.
5. After you load the Analysis ToolPak, the **Data Analysis** command is available in the **Analysis** group on the **Data** tab.
6. If you have trouble see: <http://office.microsoft.com/en-us/excel/HP100215691033.aspx>:

**Excel 2003:** The Data Analysis add-in should appear in the Tools menu. If not then

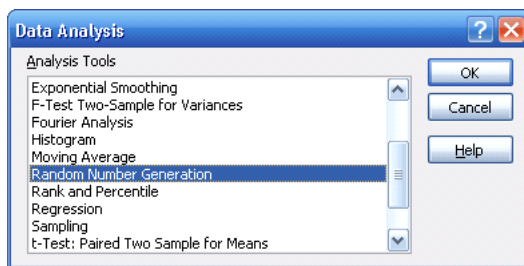
1. On the **Tools** menu, click **Add-Ins**.
2. In the **Add-Ins available** box, select the check box next to **Analysis Toolpak**, and then click **OK**.  
**Tip** If **Analysis Toolpak** is not listed, click **Browse** to locate it.
3. If you see a message that tells you the Analysis Toolpak is not currently installed on your computer, click **Yes** to install it.
4. Click **Tools** on the menu bar. When you load the Analysis Toolpak, the **Data Analysis** command is added to the **Tools** menu.
5. If you have trouble see <http://office.microsoft.com/en-us/excel/HP01127241033.aspx>

## Creating Random Numbers in Excel 2007

1. Got to the Data tab and select Data Analysis.



2. Select Random Number Generator or Sampling

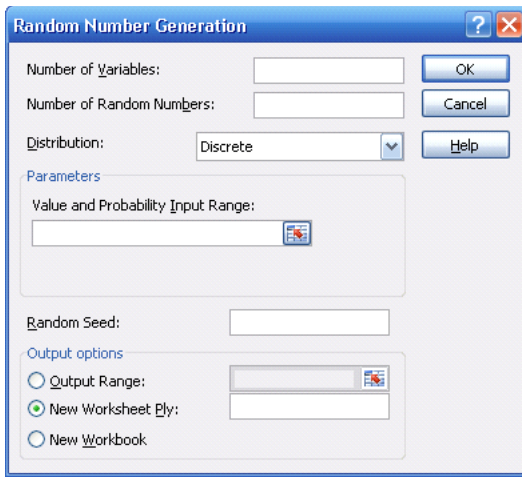


3. Enter in your information into the correct lines.

## Random Number Generator

The Random Number Generation analysis tool fills a range with independent random numbers drawn from one of several distributions. You can characterize subjects in a population with a probability distribution.

For example, you might use a normal distribution to characterize the population of individuals' heights, or you might use a Bernoulli distribution of two possible outcomes to characterize the population of coin-flip results.

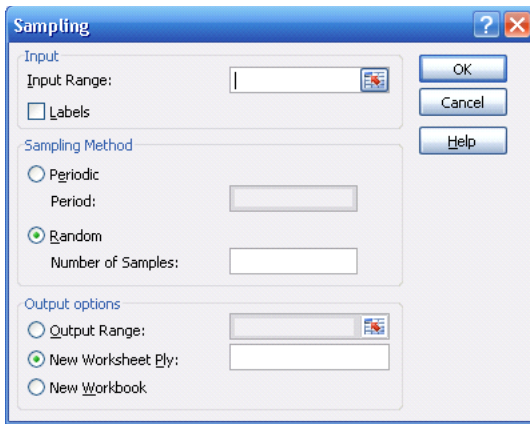


- **Number of Variables** - Enter the number of columns of values you want in the output table. If you do not enter a number, Microsoft Excel fills all columns in the output range you specify.
- **Number of Random Numbers** - Enter the number of data points you want to see. Each data point appears in a row of the output table. If you do not enter a number, Excel fills all rows in the output range you specify.
- **Distribution** - Click the distribution method you want to use to create random values.
  - **Uniform** - Characterized by lower and upper bounds. Variables are drawn with equal probability from all values in the range. A common application uses a uniform distribution in the range 0...1.
  - **Normal** - Characterized by a mean and a standard deviation. A common application uses a mean of 0 and a standard deviation of 1 for the standard normal distribution.
  - **Bernoulli** - Characterized by a probability of success (p value) on a given trial. Bernoulli random variables have the value 0 or 1. For example, you can draw a uniform random variable in the range 0...1. If the variable is less than or equal to the probability of success, the Bernoulli random variable is assigned the value 1; otherwise, it is assigned the value 0.
  - **Binomial** - Characterized by a probability of success (p value) for a number of trials. For example, you can generate number-of-trials Bernoulli random variables, the sum of which is a binomial random variable.
  - **Poisson** - Characterized by a value lambda, equal to 1/mean. Poisson distribution is often used to characterize the number of events that occur per unit of time — for example, the average rate at which cars arrive at a toll plaza.
  - **Patterned** - Characterized by a lower and upper bound, a step, repetition rate for values, and repetition rate for the sequence.
  - **Discrete** - Characterized by a value and the associated probability range. The range must contain two columns: The left column contains values, and the right column contains probabilities associated with the value in that row. The sum of the probabilities must be 1.
- **Parameters** - Enter values to characterize the distribution selected.
- **Random Seed** - Enter an optional value from which to generate random numbers. You can reuse this value later to produce the same random numbers.
- **Output Range** - Enter the reference for the upper-left cell of the output table. Excel automatically determines the size of the output area and displays a message if the output table will replace existing data.
- **New Worksheet Ply** - Click to insert a new worksheet in the current workbook and paste the results starting at cell A1 of the new worksheet. To name the new worksheet, type a name in the box.
- **New Workbook** - Click to create a new workbook and paste the results on a new worksheet in the new workbook.

## Sampling

The Sampling analysis tool creates a sample from a population by treating the input range as a population. When the population is too large to process or chart, you can use a representative sample. You can also create a sample that contains only values from a particular part of a cycle if you believe that the input data is periodic.

For example, if the input range contains quarterly sales figures, sampling with a periodic rate of four places values from the same quarter in the output range.



- **Input Range** - Enter the references for the range of data that contains the population of values you want to sample. Microsoft Excel draws samples from the first column, then the second column, and so on.
- **Labels** - Select if the first row or column of your input range contains labels. Clear if your input range has no labels; Excel generates appropriate data labels for the output table.
- **Sampling Method** - Click Periodic or Random to indicate the sampling interval you want.
- **Period** - Enter the periodic interval at which you want sampling to take place. The period-th value in the input range and every period-th value thereafter is copied to the output column. Sampling stops when the end of the input range is reached.
- **Number of Samples** - Enter the number of random values you want in the output column. Each value is drawn from a random position in the input range, and any number can be selected more than once.
- **Output Range** - Enter the reference for the upper-left cell of the output table. Data is written in a single column below the cell. If you select Periodic, the number of values in the output table is equal to the number of values in the input range, divided by the sampling rate. If you select Random, the number of values in the output table is equal to the number of samples.
- **New Worksheet Ply** - Click to insert a new worksheet in the current workbook and paste the results starting at cell A1 of the new worksheet. To name the new worksheet, type a name in the box.
- **New Workbook** - Click to create a new workbook and paste the results on a new worksheet in the new workbook.