YASAI-UM 3.0

Description

YASAI[1] is a free Excel Add-in for performing Monte Carlo Simulations in Excel. It was developed at Rutgers University and is available from its website YASAI Simulation. Instructions for using YASAI are included in the Guide to YASAI at the YASAI website. The YASAI-UM setup program (described below) creates a link to the online Guide in the Windows Start menu. This document includes important excerpts from the online Guide to YASAI for easy reference.

YASAI-UM is a special distribution of YASAI that includes two Excel Add-ins that are named **yasai.xla** and **yasaium.xla**.

- **yasai.xla** is the same as the YASAI Excel Add-in version 3.0 from the YASAI download webpage at Rutgers.
- **yasaium.xla** is an Excel Add-in that extends **yasai.xla** and provides additional capabilities, including the functions needed for the simulation of probabilistic project scheduling networks (i.e., for simulating PERT networks). This extra functionality is described below.

Yasai-UM Installation on Excel for Windows

Under Windows, Yasai-UM has its own installation program named:

Yasai-UM_setup_3.0.exe

To install the two UM-YASAI Add-ins on your Windows PC, exit from Excel (if it is currently running) and run the above Yasai-UM setup program (it does not require administrator privileges):

This setup program will install both **yasai.xla** and **yasaium.xla** on your Windows PC and will also register them with Excel so they can be used immediately. The setup program also installs Internet links to the "YASAI" website and to the "Guide to YASAI" webpage at Rutgers, and installs example/demo files, and instructions. These files are accessible through links from the folder "YASAI-UM" that is added to the Windows Start menu.

Alternatively, you can do a manual installation of Yasai-UM as described at the end of this document.

Yasai-UM Installation on Excel for Macintosh

With a manual installation, YASAI V3.0 can also be used with Macintosh Excel 2011 and 2016.

YASAI functionality is largely identical between Excel for Windows and Excel for the Macintosh. A known exception is that *Histogram charts* are not available on Mac Excel 2011 due to an Excel bug (*cumulative charts* are available).

[1] "YASAI" is an acronym for "Yet Another Simulation Add-In". The word "yasai" means "vegetables" in Japanese.

YASAI Functions for Generating Random Variates

YASAI provides Excel functions that return random variates that follow specified distributions. These functions generally return a different value each time they are called, depending on their arguments.

GENUNIFORM (*a*, *b*): Both arguments are numbers. Normally, it is expected that $a < b$. If so, a random number uniformly distributed over the interval $[a, b)$ — that is, x such that $a \leq x < b$ — is returned. If $a = b$, then the value a (or equivalently b) is returned. If $a > b$, an error value is returned.

GENTABLE (*V*, *P*): Models Probability Mass Functions. The arguments, *V* and *P*, are blocks of cells or lists (for example, "{1,3,7}") having the same number of cells. Essentially, the function returns each value in *V* with the probability specified by the corresponding element in *P*. If the two arguments have the same number of cells but differing numbers of rows and columns, the correspondence is determined by scanning first across the first row, then across the second row, and so forth. Non-numeric entries in *P* are treated as if they were zero. If the two arguments do not have the same number of cells, or if *P* contains any negative numbers, or if *P* contains only zeroes, an error value is returned. If the values in *P* do not sum to 1, they are rescaled proportionally so that they do. For example,

GENTABLE({1,2,3},{.2,.5,.3})

returns 1 with probability 0.2, 2 with probability 0.5, and 3 with probability 0.3.

GENTRIANGULAR (*a, b, c*): Returns a value from a triangular distribution with minimum *a*, mode *b*, and maximum *c*. The arguments must be numbers such that $a \leq b \leq c$, or an error value is returned.

GENNORMAL (*m*, *s*): Generates a random variate from a normal distribution with mean *m* and standard deviation *s*. Both arguments *m* and *s* are numbers. If *s* is zero, the function returns the value *m*. If $s < 0$, the function returns an error value.

GENLOGNORMAL (*m*, *s*): Generates a lognormal random variate, and is equivalent to exp(GenNormal(*m*, *s*)). The arguments *m* and *s* are the mean and the standard deviation for the underlying normal variate and the restrictions on the arguments are the same as for GENNORMAL.

GENEXPON (*a*): The argument must be a positive number, or an error value is returned. If so, the return value is randomly chosen from an exponential distribution with mean value $1/a$.

GENPOISSON (*m*): The argument *m* is a nonnegative number. A negative argument causes an error value to be returned. A zero argument causes zero to be returned. Otherwise, the return value is randomly chosen from a Poisson distribution with mean value *m*. The implementation is efficient even when m is large.

GENBINOMIAL (*n*, *p*): The first argument *n* must be a nonnegative integer, and the second argument *p* must be a number in the range [0, 1]. Otherwise, an error value is returned. If these conditions are met, then the return value is an integer drawn randomly from a binomial distribution with *n* trials and probability *p* of success at each trial. Note that if $n = 0$, then the return value is 0. The implementation is efficient even when n is large.

GENGEOMETRIC (*p*): Returns a geometric random variate which is the number of Bernoulli experiments to the first success, when each experiment has the same probability of success *p*. The outcome of such an experiment can be modeled as GENBINOMIAL(1,*p*) which returns either 1 (success) or 0 (failure).The value of *p* must be greater than 0, and less than or equal to 1, or an error value is returned.

Additional Functions Provided by the Add-in yasaium.xla

1. Module modYASAIw

This module provides the additional procedures and functions that are included in a special fork of YASAI called "YASAIw" that was developed at:

- Department of Ecology, State of Washington: https://ecology.wa.gov/Research-Data/Data-resources/Models-spreadsheets/Modeling-the-environment/Models-tools-for-TMDLs

Instructions for using YASAIw are available at the above website. For easy reference, some of the functions provided by YASAIw are described below:

GenLognormal (m, s): Returns a value from a lognormal distribution with mean *m* and standard deviation *s*. (This function is different from GenLogNormal (*m*, *s*), where *m* and *s* are the mean the standard deviation for the corresponding Normal distribution.)

GenCFD (range): Returns a value that is randomly selected from a cumulative frequency distribution (CFD) that is specified in a contiguous two column range of Worksheet cells, where the first column of the range contains the frequency values between 0 and 1 (which must begin with 0 and end with 1 and be sorted from 0 to 1), and the second column of the range specifies the cumulative distribution values for each frequency. For example, if cells AZ26:BA36 are cells that contain the data for a CFD as follows (e.g. min=1490, 5%tile=1540, etc.), then AZ26:BA36 would be entered as the range argument:

0	1490
0.05	1540
0.1	1590
0.15	1622.5
0.25	1682.5
0.5	1820
0.75	1962.5
0.85	2042.5
0.9	2105
0.95	2190
1	2275

GenLimitNormal (m, s, min, max, optional method): Returns a value with an underlying normal distribution with mean *m* and standard deviation *s* that is bounded by the *min* and *max*. The '*method*' argument is optional and allows for calculation using one of the following two methods:

- *Method*=0 (or missing) samples according to the underlying distribution until a sampled value is within the bounds of the *min* and *max*. This method changes the percentiles relative to the underlying distribution and takes the density of the underlying distribution conditioned on being within the truncation bounds.
- *Method*=1 samples from an unbounded distribution and replaces any values that are above or below the *max* or *min* with the *max* or *min*. This method preserves the same percentiles as the unbounded distribution within the bounds of the *min* and *max*.

GenLimitLognormal (m, s, min, max, optional method): Returns a value with an underlying lognormal distribution with mean *m* and standard deviation *s* that is bounded by the *min* and *max*. The optional '*method*' is the same as described for GenLimitNormal.

GenBetaPERT (min, mode, max, optional weight): Returns a value with a PERT-beta distribution with a minimum (*min*), most likely value (*mode*), and maximum (*max*). The optional parameter (*weight*) is a weighting factor that is used to estimate the mean value (*weight*=4 is assumed if weight is not specified). As the *weight* parameter is increased, the estimated mean value approaches the most likely value (*mode*). The PERT-beta is useful for modeling expert opinion in which a variable is bounded with known or estimated bounds (*min* and *max*) and for which a "most likely value" is known or can be estimated. The PERT-beta distribution is a smoother curved alternative to a triangular distribution.
The PERT-beta distribution βp(min, mode, max, weight) is related to the standard beta distribution β(α1, α2) through the following relationships (EPA, 2007; Vose, 2009):

- βp(min, mode, max, weight) = min + (max – min) * β(α1, α2)
- α1 = ((μ - min) * (2*mode – min – max)) / ((mode - μ) * (max – min))
- α2 = (α1 * (max - μ)) / (μ - min)
- μ(mean) = (min + weight*mode + max) / (weight + 2)

The mean value (μ) of the PERT-beta distribution is assumed using the equation above in order to estimate the appropriate shape parameters (α1 and α2) of the corresponding beta distribution. Unlike the triangular distribution, the mean value (μ) of the PERT-beta distribution is more influenced by the most likely value (mode) depending on the value of the weight parameter. The weight parameter influences the spread of the distribution around the most likely value, with larger values of weight resulting in narrower spread. A default weight = 4 is suggested based on the standard assumption for PERT networks that the best estimate for the duration of a task = (min + 4*mode + max)/6 (Vose, 2009).

2. Module modPERT

This module provides functions to support probabilistic project scheduling (i.e., for simulating PERT networks) and was developed by:

- Professor Photios G. Ioannou at the University of Michigan.
  http://www.ioannou.org

Module modPERT provides three functions that return the inverse of the following cumulative distributions (note that these are deterministic functions; they are not random number generators):

BetaInvPERT(*probability*, a, M, b): Returns the value "*t*", such that P[*T*<=*t*]=*probability*.
The random variable *T* follows a Beta distribution defined by the three-point PERT estimates **a, M, b,** where
- a = minimum possible value of *T*
- M = most likely value of *T* = mode of *T*
- b = maximum possible value of *T*

BetaInvPG(*probability*, p5, M, p95): Returns the value "*t*", such that P[*T*<=*t*]=*probability*.
The random variable *T* follows a Beta distribution defined by the Perry & Greig estimates **p5, M, p95,** where
- p5 = 5th percentile of *T*
- M = most likely value of *T* = mode of *T*
- p95 = 95th percentile of *T*

TriangInv(*probability*, a, M, b): Returns the value "*t*", such that P[*T*<=*t*]=*probability*.
The random variable *T* follows a Triangular distribution defined by **a, M, b,** where
- a = minimum possible value of *T*
- M = most likely value of *T* = mode of *T*
- b = maximum possible value of *T*

Generating Random Variates for Activity Durations

Monte Carlo simulation of probabilistic project scheduling networks (PERT) in Excel requires the generation of random numbers to represent the random durations of activities. The above functions can be used for this purpose by substituting "***probability***" with the Excel function **RAND()**. This technique is called the "*Inverse Cumulative Method*" for generating random numbers. For example, to generate activity durations from a Beta distribution, or a Triangular distribution use one of the following:

- Random Activity Duration = BetaInvPERT(RAND(), a, M, b)
- Random Activity Duration = BetaInvPG(RAND(), p5, M, p95)
- Random Activity Duration = TriangInv(RAND(), a, M, b)

For more details, and a complete example, see the Excel file that includes the YASAI model for the "Highway Project" (this Excel file is included with Yasai-UM).

Specifying Scenarios in YASAI

In YASAI, decision variables are called *parameters*. When a model defines multiple *parameters*, YASAI obtains a separate statistical sample for each possible combination of all the parameter values where it records the values of all the output variables.

In many applications of simulation, we need to evaluate a number of similar models that vary only in some selected decision parameters. In such cases, the objective is to select the parameter combination that gives the highest expected profit or lowest expected cost, perhaps subject to some acceptable level of risk quantified by other some other set of outputs. We call "each tested set" of "parameter values" a *scenario*. Usually, we seek to find the best scenario.

Use the following function to specify each parameter, as well as the combinations of parameter values that YASAI should consider:

SIMPARAMETER(*L, name, group*):

- The first argument, ***L***, is "a block of cells" or "a list" specifying the possible parameter values.
- The *name* argument is a character string describing the parameter and is used only in the output reports. If it is omitted, its value is taken from the cell containing the formula -- for example, if the SIMPARAMETER function is in cell C12, the default name of the parameter is "C12".
- The *group* argument is optional and defaults to 1 if omitted. It may be any whole number between 1 and 20 (values outside this range produce an error value). For simplicity, all parameters within the same group should have the same number of values in *L*.

Ordinarily, YASAI considers all possible combinations of parameter values for parameters in different groups. The values of parameters that belong to the same group, however, vary together in "lock step" with each another. For example, suppose that a model has the following SIMPARAMETER functions:

- SIMPARAMETER({1, 2, 3}, "Fred", 1)
- SIMPARAMETER({100, 150, 200}, "George", 1)
- SIMPARAMETER({1000, 2000}, "Amy", 2)

Then YASAI will test the following 3*2=6 parameter combinations:

Fred=1	George=100	Amy=1000	Fred=1	George=100	Amy=2000
Fred=2	George=150	Amy=1000	Fred=2	George=150	Amy=2000
Fred=3	George=200	Amy=1000	Fred=3	George=200	Amy=2000

Older versions of YASAI provided a different function, called PARAMETER, for specifying parameters. This function was similar to SIMPARAMETER, but it made testing combinations of parameter values from multiple lists more complicated. You can find a description of the PARAMETER function in earlier versions of the user guide.

It is recommended that you do not mix SIMPARAMETER and PARAMETER specifications in the same model. If you do, however, all parameters specified with PARAMETER are treated as being in group 1.

Specifying Output

To specify model outputs in YASAI use the function:

SIMOUTPUT(*x*, *name*)

This method is simple and effective. Including the SIMOUTPUT function anywhere in the spreadsheet specifies that output statistics should be tabulated for the model quantity ***x*** (where ***x*** is usually a formula that returns different values during the simulation). The argument *"name"* is a string enclosed in double quotes that specifies *what the output should be called* in the *simulation output report*.

When working interactively with a spreadsheet, the SIMOUTPUT function always returns the current value of ***x*** (which should change every time we press F9 to recalculate the worksheet)—it does nothing more. However, when a simulation is running, SIMOUTPUT records all the values of ***x*** and saves them for later analysis, as described below.

Example-1: A cell containing the formula below defines an output called "Profit" whose value is B5-B6.

=SimOutput (B5-B6, "Profit")

Example-2: Formulas can combine SIMOUTPUT functions in arbitrary ways. For example, the formula below records simultaneously the value of (B5-B6) as "Profit", but also records the value of B5 as "Revenue", and the value of B6 as "Cost".

=SimOutput (SimOutput(B5, "Revenue") – SimOutput (B6, "Cost"), "Profit")

Example-3: Suppose cell C6 contains a Poisson random variable whose mean value is specified by cell C3. While C6's main purpose is to serve as an input to other formulas in the model, we would like the simulation output to contain statistics about it. To ensure the output contains this information, we just change the formula in C6 as follows:

=GenPoisson(C3)	=SimOutput (GenPoisson(C3), "Poisson Input Variable")

Example-4: Creative use of conditional formulas can afford even greater flexibility. Suppose that cell D12 contains net cash flow. The formula below will create two outputs. The first, "Profit", will contain statistics on the net cash flow in D12, but tabulated only over those instances where it was nonnegative. The second, "Loss", will tabulate the absolute value of net cash flow, but only over those instances where net cash flow was negative.

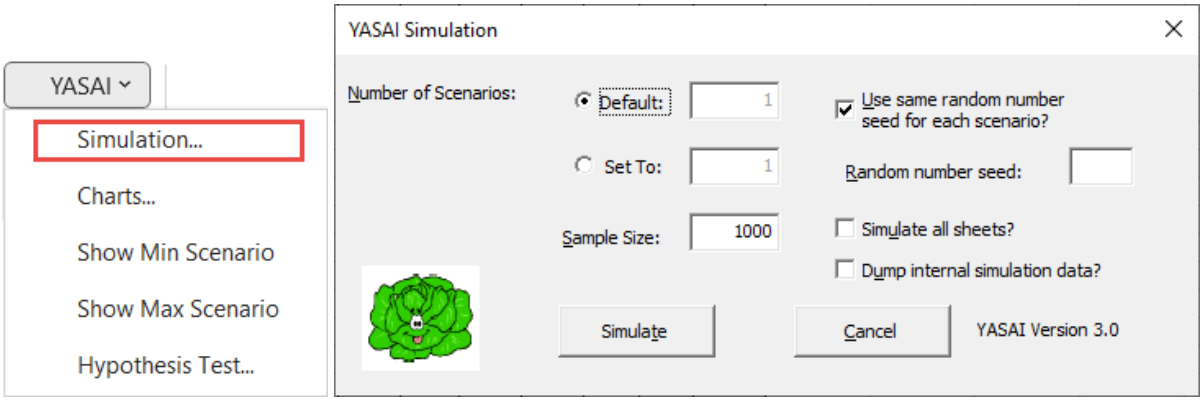
=SimOutput (ABS(D12), IF(D12>=0, "Profit", "Loss"))

Moreover, if multiple SIMOUTPUT expressions in a model have the same *name* argument, their values are pooled together in the simulation output.

YASAI's method for specifying *simulation outputs* using *formulas* has two important advantages:
- The output formulas can be examined and modified easily.
- The output formulas are saved as part of the same Excel file as the model and thus cannot be *lost*.

Running the Simulation

Once we have built our model, specified scenarios (if any), and specified simulation outputs using the function SIMOUTPUT, we can run the simulation. To do so, click on the “Add-ins” tab of the Excel ribbon and select "YASAI > Simulation". This will bring up the following dialog box on the right.



Number of scenarios: YASAI analyzes the spreadsheet to determine how many scenarios may be needed. It places this number in the default number of scenarios box. If this number is satisfactory, click on the "Default" button. If you would like a different number of scenarios, click on the "set to" button and enter the number of scenarios you want.

The "Sample Size" box is the number of times YASAI will recalculate your model for each scenario. It defaults to 1000, but we can enter any positive whole number.

Optionally, we may indicate a fixed random-number-seed to use, and whether each scenario should use the same seed (using the same seed for each scenario is good practice and is the default).

Press "Simulate" to start the simulation, or "Cancel" to return to Excel.

Unless “Simulate all sheets” is checked, the simulation involves only the current sheet of the current workbook. If the number of scenarios" is *N*, and the "Sample Size" is *S*, then YASAI will recalculate the current sheet *N***S* times. These recalculations are divided into *N* blocks of *S* recalculations, each block constituting a scenario. Output data are collected separately for each scenario, and the values returned by any SIMPARAMETER functions in the sheet will vary from scenario to scenario, as described above.

A "progress" display indicates how quickly the simulation is progressing. When the simulation is over, there will be a short delay while the outputs are processed. The output report is automatically placed in a new sheet named "Simulation Output *n*", which YASAI inserts into the current workbook. YASAI makes this report the current sheet and then returns control to Excel. For each output-scenario combination, the report contains the mean, standard deviation, minimum, maximum, and percentiles in 5% intervals. Currently there are no graphics or other output data, although improvements are planned for later versions.

Aborting a Simulation

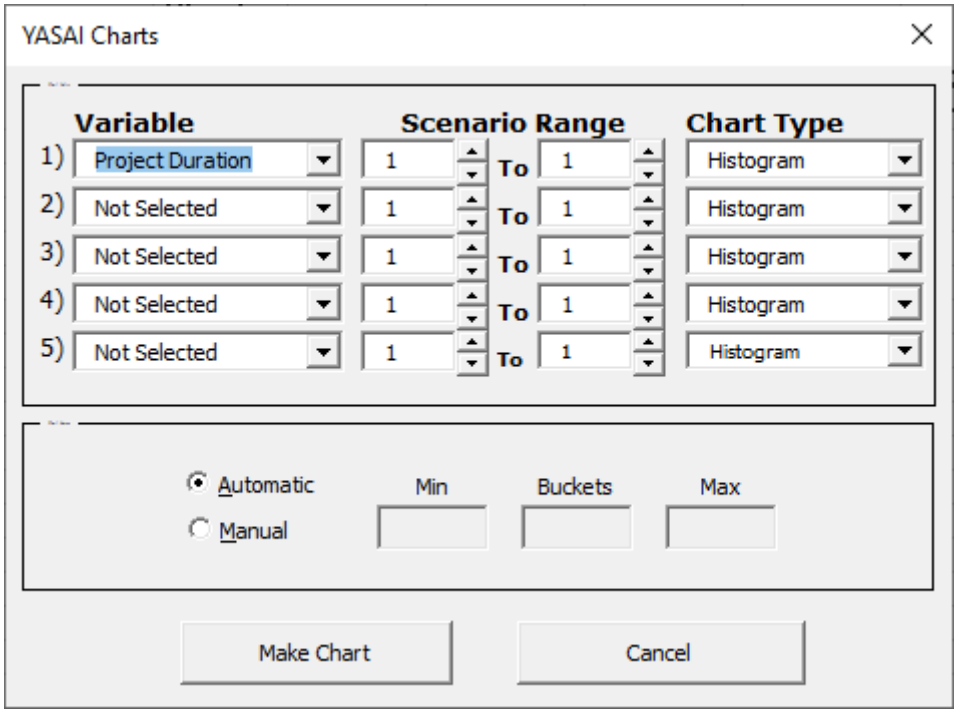
You can abort a YASAI simulation while it is running by clicking on the "Abort" button, or simply by pressing the escape key on your keyboard. It may take up to 5 seconds for the simulation to abort.

Testing a Model Interactively

To test a model interactively, simply press the **F9** key. Excel will perform a single recalculation, drawing new values for all the random variate generation functions. New values will also be generated for each SIMPARAMETER function call.

Charting Simulation Output

YASAI can chart simulation outputs. You must run a simulation before trying to produce charts. To make a chart, select "YASAI Charts..." from the Add-ins tab of the ribbon, which produces a charting window:



You may chart up to five blocks of outputs, each corresponding to one row in the window. The first column selects which variables to chart. The second and third columns allow you to specify a range of scenarios for the selected variable, for example scenarios 1 through 5; a graph for each scenario in the range will appear in the output. The last column selects the kind of graph desired for the block of variables: "Histogram" is a standard bar chart, and "Cumulative plot" produces empirical cumulative distributions. You may mix the two kinds of graphs on a single chart. With the "automatic" button set under "Chart Range" YASAI chooses the horizontal axis range and subdivisions to attempt to produce an attractive chart. "Manual" lets you specify a range from "Min" to "Max", with "Buckets" subdivisions.

YASAI charts are regular Excel charts. Once they have been created, they can be modified as needed. The chart and its associate data each appear in separate sheets. You may delete them if they are no longer needed. YASAI charts are "static" — they are based on the simulation immediately preceding their creation; if you run another simulation they will not update. You must run a simulation again and make new charts if you want your charts to reflect a change to your simulation model.
Note: Due to internal limitations in Excel, cumulative graphs are not possible for outputs with more than 32,760 observations. Histograms are possible for any sample size.

Yasai-UM manual installation

The files for both Yasai-UM Add-ins, **yasai.xla** and **yasaium.xla**, should be placed in the following directory:

%APPDATA%\Microsoft\AddIns

The easiest way to access this directory is to run the Windows File Explorer and copy-paste (or type) the line above exactly as shown (including the environment variable %APPDATA%) into the directory destination.

The Windows File Explorer will automatically replace the Windows *environment variable* %APPDATA% with the proper string (for your account) and will take you to a directory like the one shown below (which includes the current user's "*username*"):

C:\Users*username*\AppData\Roaming\Microsoft\AddIns

To complete the installation, you must now run Excel and register/activate the two Yasai-UM Add-ins, **yasai.xla** and **yasaium.xla**, from inside Excel. For Excel 2016, follow the steps below:

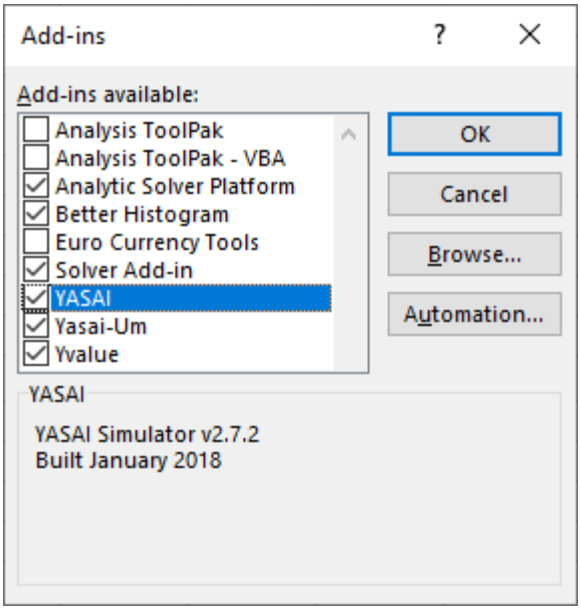
Run Excel and in a blank worksheet, select from the ribbon:

File > Options> Add-ins

At the bottom of the dialog box select:

"Manage *Excel Add-ins*" and then click on "Go…"

To register/activate the add-ins **yasai.xla** and **yasaium.xla** place checkmarks next to their names as shown below and click on OK.



The Yasai-UM setup program (described above) performs these steps (and more) automatically.

Summary of Functions that Generate Random Variates

Yasai Gen Functions	**YasaiW Gen Functions**	**EasyFit Gen Functions** (1)
GenBinomial GenExpon GenGeometric GenLogNormal GenNormal GenPoisson GenTable GenTriangular GenUniform Auxiliary Functions (do not use directly) GenBinomialWaitingTime GenBinomPage533	GenBetaPERT GenCFD GenLognormal GenLimitLognormal GenLimitNormal Correlated variables (Not described above) GenCFDX GenCFDY GenLimitLognormalX GenLimitLognormalY GenLimitNormalX GenLimitNormalY GenLognormalX GenLognormalY GenNormalX GenNormalY Auxiliary Functions (do not use directly) GenNormal2 GenZ	Not to be confused with YASAI Functions GenExtremeCdf GenExtremeHaz GenExtremeInv GenExtremeMean GenExtremePdf GenExtremeRand GenExtremeStdev GenExtremeVar GenGammaCdf GenGammaHaz GenGammaInv GenGammaMean GenGammaPdf GenGammaRand GenGammaStdev GenGammaVar GenLogisticCdf GenLogisticHaz GenLogisticInv GenLogisticPdf GenLogisticRand GenParetoCdf GenParetoHaz GenParetoInv GenParetoMean GenParetoPdf GenParetoRand GenParetoStdev GenParetoVar
Module modPERT Functions		
BetaInvPERT(probability, a, M, b) BetaInvPG(probability, p5, M, p95) TriangInv(probability, a, M, b)		

(1) The EasyFit functions for generating random variates in Excel are provided by the EasyFit add-in for Excel which is automatically installed along with the commercial version of EasyFit. Unfortunately, this add-in is only available as a 32-bit DLL that works only with the 32-bit version of Excel. It does not work with the 64-bit version of Excel. And, there is no 64-bit version of the DLL available.