rocscience

Probabilistic Analysis Tutorial

1. Introduction

This tutorial will familiarize the user with the basic probabilistic analysis capabilities of Slide2. It will demonstrate how quickly and easily a probabilistic slope stability analysis can be performed with Slide2.

The finished product of this tutorial can be found in the Tutorial 08 Probabilistic Analysis.slmd data file. All tutorial files installed with Slide2 can be accessed by selecting **File > Recent Folders > Tutorials Folder** from the Slide2 main menu.

2. Model

This tutorial will be based on the same model used for Tutorial 1, so let's first read in the Tutorial 1 file.

Select **File > Recent Folders > Tutorials Folder from the Slide2** main menu, and open the *Tutorial 01 Quick Start.slmd* file.

PROJECT SETTINGS

To carry out a Probabilistic Analysis with Slide2, the first thing that must be done is to select the Probabilistic Analysis option in the Project Settings dialog.

Select **Analysis > Project Settings**

In the Project Settings dialog, select the **Statistics** page, and select the **Probabilistic Analysis** checkbox.

Notice the Spatial Variability Analysis checkbox. For more information on running a spatially variable analysis, see Tutorials 33 and 34.

Select **OK**.

GLOBAL MINIMUM ANALYSIS

Note that we are using the default Probabilistic Analysis options:

- Sampling Method = Latin Hypercube
- Number of Samples = 1000

- Analysis Type = Global Minimum

When the Analysis Type = Global Minimum, this means that the Probabilistic Analysis is carried out on the Global Minimum slip surface located by the regular (deterministic) slope stability analysis.

The safety factor will be re-computed N times (where N = Number of Samples) for the Global Minimum slip surface, using a different set of randomly generated input variables for each analysis.

Notice that a Statistics menu is now available, which allows you to define almost any model input parameter as a random variable.

DEFINING RANDOM VARIABLES

In order to carry out a Probabilistic Analysis, at least one of your model input parameters must be defined as a Random Variable. Random variables are defined using the options in the Statistics menu.

For this tutorial, we will define the following material properties as Random Variables:

- Cohesion
- Friction Angle
- Unit Weight

This is easily done with the Material Statistics dialog.

Select **Statistics > Materials**

You will see the Material Statistics dialog. Notice that “Soil 1” is selected on the left side. We are now going to define Random Variables for this material. This can be done with either the Add or the Edit options, in the Material Statistics dialog. Let’s use the Add option.

Select the Add button in the Material Statistics dialog.

When using the Add option, you will see a series of three dialogs, in a “wizard” format, which allow you to quickly select the material properties that you wish to define as Random Variables.

The first dialog allows you to select the material properties that you would like to define as Random Variables.

Select the checkboxes for Cohesion, Phi and Unit Weight. Select the Next button. The second dialog allows you to select a Statistical Distribution for the Random Variables. We will be using the default (Normal Distribution), so just select the Finish button.

You will be returned to the Material Statistics dialog, which should now appear as follows:

In the Material Statistics dialog, the material properties which you selected as Random Variables, now appear in the dialog in a spreadsheet format. This allows you to easily define the statistical distribution for each random variable.

In order to complete the process of defining the Random Variables, we must enter:

- Standard Deviation
- Minimum and Maximum values

for each variable, in order to define the statistical distribution of each random variable.

Enter the values of Standard Deviation, Relative Minimum and Relative Maximum for each variable, as shown below. When you are finished, select OK.

Variable #	Standard Deviation	Relative Minimum	Relative Maximum
1	1	3	3
2	3	9	9
3	0.5	1.5	1.5

Note

- The Minimum and Maximum values are specified as RELATIVE values (i.e. distances from the MEAN value), rather than as absolute values because this simplifies data input.
- For a NORMAL distribution, 99.7 % of all samples should fall within 3 standard deviations of the mean value. Therefore it is recommended that the Relative Minimum and Relative Maximum values are equal to at least 3 times the standard deviation, to ensure that a complete (non-truncated) NORMAL distribution is defined.
- For more information about Statistical Distributions, please see the Probabilistic Analysis section of the Slide2 Help system.

That's all we need to do. We have defined 3 Random Variables (cohesion, friction angle and unit weight) with Normal distributions.

We can now run the Probabilistic Analysis.

3. Compute

First, let's save the file with a new file name: *Tutorial 08.slmd.*

Select **File > Save As**

Use the Save As dialog to save the file. Now select Compute.

Select **Analysis > Compute**

Note

- When you run a Probabilistic Analysis with Slide2, the regular (deterministic) analysis is always computed first.
- The Probabilistic Analysis automatically follows. The progress of the analysis is indicated in the Compute dialog.

4. Interpret

To view the results of the analysis:

Select **Analysis > Interpret**

This will start the Slide2 Interpret program. You should see the following figure.

The primary results of the probabilistic analysis, are displayed beside the slip center of the deterministic global minimum slip surface. Remember that when the Probabilistic Analysis Type = Global Minimum, the Probabilistic Analysis is only carried out on this surface.

This includes the following:

• FS (mean) – the mean safety factor

• PF – the probability of failure

• RI – the Reliability Index

Summary of results after probabilistic analysis.

These results are discussed below.

DETERMINISTIC SAFETY FACTOR

The Deterministic Safety Factor, FS (deterministic), is the safety factor calculated for the Global Minimum slip surface, from the regular (non-probabilistic) slope stability analysis.

This is the same safety factor that you would see if you were only running a regular (deterministic) analysis, and were NOT running a Probabilistic Analysis.

The Deterministic Safety Factor is the value of safety factor when all input parameters are exactly equal to their mean values.

MEAN SAFETY FACTOR

The Mean Safety Factor is the mean (average) safety factor, obtained from the Probabilistic Analysis. It is simply the average safety factor, of all of the safety factors calculated for the Global Minimum slip surface.

In general, the Mean Safety Factor should be close to the value of the deterministic safety factor, FS (deterministic). For a sufficiently large number of samples, the two values should be nearly equal.

PROBABILITY OF FAILURE

The Probability of Failure is simply equal to the number of analyses with safety factor less than 1, divided by the total Number of Samples.

For this example, PF = 10.2 %, which means that 102 out of 1000 samples produced a safety factor less than 1.

RELIABILITY INDEX

The Reliability Index is another commonly used measure of slope stability, after a probabilistic analysis.

The Reliability Index is an indication of the number of standard deviations which separate the Mean Safety Factor from the critical safety factor (= 1).

The Reliability Index can be calculated assuming either a Normal or Lognormal distribution of the safety factor results. The actual best fit distribution is listed in the Report Generator and indicates which value of RI is more appropriate for the data.

RI (Normal)

If it is assumed that the safety factors are Normally distributed, then Equation 2 is used to calculate the Reliability Index.

A Reliability Index of at least 3 is usually recommended, as a minimal assurance of a safe slope design. For this example, RI = 1.2, which indicates an unsatisfactory level of safety for the slope.

RI (Lognormal)

If it is assumed that the safety factors are the best fit by a Lognormal distribution, then Equation 3 is used to calculate the Reliability Index.

For more information about the Reliability Index, see the Slide2 Help system.

HISTOGRAM PLOTS

Histogram plots allow you to view:

• The distribution of samples generated for the input data random variable(s).

• The distribution of safety factors calculated by the probabilistic analysis.

To generate a Histogram plot, select the Histogram Plot option from the toolbar or the Statistics menu.

Select **Statistics > Histogram Plot**

You will see the Histogram Plot dialog.

Let's first view a histogram of Safety Factor. Set the Data to Plot = Factor of Safety – Bishop Simplified. Select the Highlight Data checkbox. As the highlight criterion, select "Factor of Safety – Bishop Simplified < 1". Select the Plot button, and the Histogram will be generated.

Histogram of Safety Factor

As you can see on the histogram, the highlighted data (red bars) shows the analyses which resulted in a safety factor less than 1.

• This graphically illustrates the Probability of Failure, which is equal to the area of the histogram which is highlighted (FS < 1), divided by the total area of the histogram.

• The statistics of the highlighted data are listed on the plot. In this case, it is indicated that 103/1000 points have a safety factor less than 1. This equals 10.3%, which is the PROBABILITY OF FAILURE (for the Bishop analysis method).

In general, the Highlight data option allows you to highlight any user-defined subset of data on a histogram (or scatter plot), and obtain the statistics of the highlighted (selected) data subset.

You can display the Best Fit distribution for the safety factor data, by right-clicking on the plot and selecting Best Fit Distribution from the popup menu. The Best Fit Distribution will be displayed on the Histogram. In this case, the best fit is a Gamma Distribution, as listed at the bottom of the plot.

Let's create a plot of the Cohesion random variable. Right-click on the plot and select Change Plot Data. Set the Data to Plot = soil 1: Cohesion. Select Done.

Histogram Plot of Cohesion

This plot shows the actual random samples which were generated by the Latin Hypercube sampling of the statistical distribution which you defined for the Cohesion random variable. Notice that the data with Bishop Safety Factor < 1 is still highlighted on the plot.

Note the following information at the bottom of the plot:

• The SAMPLED statistics, are the statistics of the raw data generated by the Monte Carlo sampling of the input distribution.

• The INPUT statistics, are the parameters of the input distribution which you defined for the random variable, in the Material Statistics dialog.

In general, the SAMPLED statistics and the INPUT statistics will not be exactly equal. However, as the Number of Samples increases, the SAMPLED statistics should approach the values of the INPUT parameters.

The distribution defined by the INPUT parameters is plotted on the Histogram. The display of this curve can be turned on or off, by right-clicking on the plot and toggling the Input Distribution option.

Now right-click on the plot again, and select Change Plot Data. Change the Data to Plot to soil 1: Phi. Select Done.

Notice the data with Bishop Safety Factor < 1, highlighted on the plot. With respect to the Friction Angle random variable, it is clear that failure corresponds to the lowest friction angles which were generated by the random sampling.

CUMULATIVE PLOTS

To generate a Cumulative plot, select the Cumulative Plot option from the toolbar or the Statistics menu.

Select **Statistics > Cumulative Plot**

You will see the Cumulative Plot dialog.

Select the Data to Plot = Factor of Safety – Bishop Simplified. Select the Plot button.

Cumulative Plot of Safety Factor

A Cumulative distribution plot represents the cumulative probability that the value of a random variable will be LESS THAN OR EQUAL TO a given value.

When we are viewing a Cumulative Plot of Safety Factor, the Cumulative Probability at Safety Factor = 1, is equal to the PROBABILITY OF FAILURE.

Let’s verify this as follows.

SAMPLER OPTION

The Sampler Option on a Cumulative Plot allows you to easily determine the coordinates at any point along the Cumulative distribution curve.

1. Right-click on the Cumulative Plot, and select the Sampler > Show Sampler option.
2. You will see a dotted cross-hair line on the plot. This is the “Sampler”, and allows you to graphically obtain the coordinates of any point on the curve.

3. As you move the mouse over the cumulative curve, you will see that the Sampler follows the mouse, and continuously displays the coordinates of points on the Cumulative plot curve.
4. If you click the mouse at the point Factor of Safety = 1, notice that the Cumulative Probability = 0.103, as shown in the figure below. This means that the Probability of Failure (Bishop analysis method) = 10.3%, which is the value we noted earlier in this tutorial.

Sampler display on cumulative plot

SCATTER PLOTS

Scatter Plots allow you to plot any two random variables against each other, on the same plot. This allows you to analyze the relationships between variables.

Select the Scatter Plot option from the toolbar or the Statistics menu.

Select **Statistics > Scatter Plot**

You will see the Scatter Plot dialog. Enter the following data.

1. Set the Horizontal Axis = Soil 1: Phi (deg)

2. Set the Vertical Axis = Factor of Safety – bishop simplified

3. Select Highlight Data, and select "Factor of Safety – bishop simplified < 1".

4. Select Plot.

You should see the following plot.

Scatter Plot – Friction Angle versus Safety Factor

There is a well-defined relationship between Friction Angle and Safety Factor. Notice the parameters listed at the bottom of the plot.

• The Correlation Coefficient indicates the degree of correlation between the two variables plotted. A Correlation Coefficient close to 1 (or -1) indicates a high degree of correlation. A Correlation Coefficient close to zero indicates little or no correlation.

• The parameters Alpha and Beta, are the slope and y-intercept, respectively, of the best fit (linear) curve, to the data. This line can be seen on the plot. Its display can be toggled on or off, by right-clicking on the plot and selecting the Regression Line option.

Also, notice the highlighted data on the plot. All data points with a Safety Factor less than 1, are displayed on the Scatter Plot as an orange square, rather than a blue triangle.

Now let's plot Phi versus Cohesion on the Scatter Plot.

Right-click on the plot and select Change Plot Data. On the Vertical Axis, select soil 1: Cohesion. Select Done. The plot should look as follows:

Scatter Plot – Friction Angle versus Cohesion

This plot indicates that there is no correlation between the sampled values of Cohesion and Friction Angle. (The Correlation Coefficient, listed at the bottom of the plot, is a small number close to zero).

In reality, the Cohesion and Friction Angle of Mohr-Coulomb materials are generally correlated, such that materials with low Cohesion often have high Friction Angles, and vice versa.

In Slide2, the user can define a correlation coefficient for Cohesion and Friction Angle, so that when the samples are generated, Cohesion and Friction Angle will be correlated. This is discussed at the end of this tutorial.

CONVERGENCE PLOTS

A Convergence Plot is useful for determining whether or not your Probabilistic Analysis is converging to a final answer, or whether more samples are required.

Select the Convergence Plot option from the toolbar or the Statistics menu.

Select **Statistics > Convergence Plot**

You will see the Convergence Plot dialog. Select Probability of Failure % for Iteration Data. Select Plot.

You should see the following plot.

Convergence plot – Probability of Failure

A convergence plot should indicate that the final results of the Probabilistic Analysis, are converging to stable, final values (i.e. Probability of Failure, Mean Safety Factor etc.)

If the convergence plot indicates that you have not achieved a stable, final result, then you should increase the Number of Samples, and re-run the analysis.

Right-click on the plot and select the Final Value option from the popup menu. A horizontal line will appear on the plot, which represents the final value (in this case, Probability of Failure = 10.3%), which was calculated for the analysis.

For this model, it appears that the Probability of Failure has achieved a constant final value. To verify this, increase the Number of Samples (e.g. 2000), and re-run the analysis. This is left as an optional exercise.

5. Additional Exercises

The user is encouraged to experiment with the Probabilistic Analysis modelling and data interpretation features in Slide2. Try the following exercises.

CORRELATION COEFFICIENT (C AND PHI)

Earlier in this tutorial, we viewed a Scatter Plot of Cohesion versus Friction Angle.

Because the random sampling of these two variables, was performed entirely independently, there was no correlation between the two variables.

In reality, the Cohesion and Friction Angle of Mohr-Coulomb materials are generally correlated, such that materials with low Cohesion tend to have high Friction Angles, and vice versa.

In Slide2, you can easily define a correlation coefficient for Cohesion and Friction Angle, so that when the samples are generated, Cohesion and Friction Angle will be correlated.

This can be demonstrated as follows:

1. In the Slide2 Model program, select the Material Statistics option in the Statistics menu.

2. In the Material Statistics dialog, select the Correlation option. This will display a dialog, which allows you to define a correlation coefficient between cohesion and friction angle (this is only applicable for materials which use the Mohr-Coulomb strength type).

3. In the correlation dialog, select the Apply checkbox for “soil 1”. We will use the default correlation coefficient of –0.5. Select OK in the Correlation dialog. Select OK in the Material Statistics dialog.

4. Re-compute the analysis.

5. In the Slide2 Interpret program, create a Scatter Plot of Cohesion versus Friction Angle. You should see the following. Cohesion vs. Phi (Correlation = – 0.5)

As you can now see, Cohesion and Friction Angle are no longer independent of each other but are loosely correlated.

Note

• The actual correlation coefficient generated by the sampling, is listed at the bottom of the plot. It is not exactly equal to – 0.5, because of the relatively small number of samples (1000).

• A NEGATIVE correlation coefficient simply means that when one variable increases, the other is likely to decrease, and vice versa.

Now try the following:

1. Re-run the analysis using correlation coefficients of – 0.6, – 0.7, – 0.8, – 0.9, – 1.0. View a scatter plot of Cohesion versus Friction Angle, after each run.

2. You will see that the two variables will be increasingly correlated. When the correlation coefficient = – 1.0, the Scatter Plot will result in a straight line.

COHESION vs. PHI (CORRELATION = – 1.0)

In general, it is recommended that a correlation coefficient is defined between Cohesion and Friction Angle, for a Mohr-Coulomb material. This will generate values of Cohesion and Friction Angle, which are more likely to occur in the field.

Finally, it is interesting to note that the Probability of Failure, for this model, decreases significantly, as the correlation between cohesion and friction angle increases (i.e. closer to – 1).

This implies that the use of a correlation coefficient, and the generation of more realistic combinations of Cohesion and Phi, tends to decrease the calculated probability of failure, for this model.

SAMPLING METHOD

The default method of Random Sampling in Slide2 is known as Latin Hypercube Sampling. The other sampling methods available in Slide2 are Monte Carlo and Response Surface.

Monte Carlo

For a given number of samples, Latin Hypercube sampling results in a smoother, more uniform sampling of the probability density functions which you have defined for your random variables, compared to the Monte Carlo method. To illustrate this, do the following:

1. In the Slide2 Modeler program, select **Project Settings > Statistics**, and set the Sampling Method to Monte Carlo.
2. Re-compute the analysis.
3. View the results in Interpret, and compare with the previous (Latin Hypercube) results. In particular, plot histograms of your input random variables (Cohesion, Phi, Unit Weight).
4. Notice that the input data distributions which you defined for your input random variables are much more smoothly sampled by Latin Hypercube sampling, compared to Monte Carlo sampling.

Comparison of Monte Carlo sampling (left) and Latin Hypercube sampling (right) – Cohesion random variable – 1000 samples

Monte Carlo Sampling Comparison Charts

As you can see in the above figure, for 1000 samples, the Latin Hypercube sampling is much smoother than the Monte Carlo sampling. This is because the Latin Hypercube method is based upon "stratified" sampling, with random selection within each stratum. Typically, an analysis using 1000 samples obtained by the Latin Hypercube technique will produce comparable results to an analysis of 5000 samples using the Monte Carlo method. In general, the Latin Hypercube method allows you to achieve similar results to the Monte Carlo method, with a significantly smaller number of samples.

Response Surface

The Response Surface sampling method also uses the Latin Hypercube sampling, so that the histogram for Latin Hypercube shown above would be identical for Response Surface. However, the Response Surface sampling method uses a small number of strategically selected computations (40 computations in this case) to create a response surface of factor of safety values for various combinations of input parameters. It then predicts the factor of safety values for any combination of samples and provides an estimated probability of failure. In this case, after being trained with 40 computations, it would predict the FS value for the 1000 Latin Hypercube samples.

Since an Overall Slope probabilistic analysis can take hours, this method is advantageous in significantly cutting down computation time. It is not recommended for use with Global Minimum since this method is already fast. However, for comparison purposes, the model above has been computed with Response Surface as well, with the results displayed below.

The three input distributions are found below. As mentioned, Latin Hypercube and Response Surface use the same sampling method:

Comparison of Monte Carlo sampling (left), Latin Hypercube sampling (middle), Response Surface sampling (right) – Cohesion random variable – 1000 samples

Comparison Charts

Sampling Method	Probability of Failure	Mean FS	Num. Simulations actually computed
Monte Carlo	11.100%	1.145	1000
Latin Hypercube	10.300%	1.145	1000
Response Surface	10.400%	1.145	40

Comparison of Latin Hypercube FS (left) and Response Surface FS (right)

Comparison of Latin Hypercube FS (left) and Response Surface FS (right)

The histograms are in very good agreement, even though the FS values were predicted in the Response Surface case. Latin Hypercube found 103 points failing out of 1000 (10.3%) while Response Surface found 104 points failing (10.4%).

RANDOM NUMBER GENERATION

The sampling of the statistical distributions of your input data random variables is achieved by the generation of random numbers. You may wonder why the results in this tutorial are reproducible if they are based on random numbers?

The reason for this is that we have been using the Pseudo-Random option, in Project Settings. Pseudo-random analysis means that the same sequence of random numbers is always generated because the same “seed” value is used. This allows you to obtain reproducible results for a Probabilistic Analysis.

Try the following:

1. Select **Project Settings > Random Numbers,** and select the Random option (instead of Pseudo-Random).

2. Re-compute the analysis.

3. You will notice that each time you re-compute, the probabilistic analysis results will be different. This is because a different “seed” value is used each time. This will give a different sequence of random numbers, and therefore a different sampling of your random variables, each time you re-run the analysis.