XY-plots

Scatter plots, Regression lines and Curve fits are all referred to as “XY-plots”. Such plots are produced using data from a pair of data nodes.

DynRisk allows you to produce any combination of the three XY-plot types in the same diagram.

Moreover, you can copy curves from one plot window and paste it into another using the “Copy” and “Paste” commands in the “Edit” menu. If you do this, make sure that both plots have the same scale on both the X-axis and the Y-axis by using the “Set scale” command.

Note also that you can add “gridlines” to the plots corresponding to certain key statistics for the data nodes. Specifically, you can have gridlines indicating the location of the following statistics:

- Base value
- Mean ± st.dev.
- Fractiles

The fractiles are chosen according to the currently selected fractile set. You can change this using the “Fractiles…” command.

To display the gridlines, click anywhere in the plot window while pressing the “Command” key on the keyboard. This will bring up a popup menu from which you could select the desired statistics. To hide the gridlines, just repeat the same procedure.

Scatter plot

A “Scatter plot” is a graph consisting of a “cloud” of points reflecting the joint distribution of a pair of data nodes. Each point in the cloud represents corresponding values of the two data nodes from a single simulation. So if you make a “Scatter plot” from a simulation data file where each data node contains say 1000 values, then the plot will show a cloud of 1000 points. If the points are spread all over the diagram with no particular structure, this indicates that there is no dependence between the two variables. In particular the correlation between the two data nodes is neglectable. On the other hand if the points are clustered around a straight line, the correlation may be significant. If the points are clustered closely around some nonlinear curve, the correlation may be small. Still there appears to some sort of dependence between the two data nodes.

If the frontmost window is a main document window for a simulation data file, the “Scatter plot” command produces scatter plots for all selected pairs of data nodes. To select a pair of data nodes, select the first one whose values should appear along the X-axis in the “First selection” mode, and the second one whose values should appear along the Y-axis in the “Second selection” mode.

If the frontmost window is an “XY-plot” window created by using the “Scatter plot” command, this menu item is disabled. In this case the window title is of the form:

<node1>(x) <node2>(y).scat

where <node1> is the name of the node whose values appear along the X-axis and <node2> is the name of the node whose values appear along the Y-axis.

If the frontmost window is an “XY-plot” window created by using either the “Regression line” command or the “Curve fit” command, this menu item is changed to either “Hide Scatter plot” or “Show Scatter plot” depending on the state of the window. If the plot already contains a scatter plot, you can use the “Hide Scatter plot” command to hide this scatter. On the other hand, if the plot does not contain a scatter plot, then you can use the “Show Scatter plot” command to show it.

Regression line

A “Regression line” is a graph that represents a best linear approximation to the relation between two data nodes. To see how well this line describes the “real” relation between the data nodes, you can view the regression line in the same diagram as a scatter plot. If the points are clustered closely around the regression line, the linear approximation is good. You can also compare the regression line to a curve fit. If there is a strong linear relation between the data nodes, then the regression line and the curve fit will be similar. In such cases, you do not gain much by using a more flexible (and thus also more complex) functional relationship as offered by the curve fit.

If the frontmost window is a main document window for a simulation data file, the “Regression line” command produces regression lines for all selected pairs of data nodes. To select a pair of data nodes, select the first one whose values should appear along the X-axis in the “First selection” mode, and the second one whose values should appear along the Y-axis in the “Second selection” mode.

If the frontmost window is an “XY-plot” window created by using the “Regression line” command, this menu item is disabled. In this case the window title is of the form:

<node1>(x) <node2>(y).reg

where <node1> is the name of the node whose values appear along the X-axis and <node2> is the name of the node whose values appear along the Y-axis.

If the frontmost window is an “XY-plot” window created by using either the “Scatter plot” command or the “Curve fit” command, this menu item is changed to either “Hide regr. line” or “Show regr. line” depending on the state of the window. If the plot already contains a regression line, you can use the “Hide regr. line” command to hide this regression line. On the other hand, if the plot does not contain a regression line, then you can use the “Show regression line” command to show it.

Curve fit

A “Curve fit” is a graph that attempts to describe the functional relation between two data nodes. To see how well this curve describes the “real” relation between the data nodes, you can view the curve fit in the same diagram as a scatter plot. If the points are clustered closely around the curve fit, there is a strong functional relation between the two data nodes. You can also compare the curve fit to a regression line. If there is a strong linear relation between the data nodes, then the regression line and the curve fit will be similar. In such cases, you do not gain much by using a more flexible (and thus also more complex) functional relationship as offered by the curve fit.

DynRisk calculates the curve fit by modeling the relation between the two data nodes as a so-called Gaussian process. You can control how this is done by adjusting parameters such as “Resolution”, “Smoothness” and “Sensitivity”. To do this you select the node you want to adjust the parameter settings for and then use the “Curve fit settings…” command.

If the frontmost window is a main document window for a simulation data file, the “Curve fit” command produces curve fits for all selected pairs of data nodes. To select a pair of data nodes, select the first one whose values should appear along the X-axis in the “First selection” mode, and the second one whose values should appear along the Y-axis in the “Second selection” mode.

If the frontmost window is an “XY-plot” window created by using the “Curve

fit” command, this menu item is disabled. In this case the window title is of the form:

<node1>(x) <node2>(y).gf

where <node1> is the name of the node whose values appear along the X-axis and <node2> is the name of the node whose values appear along the Y-axis.

If the frontmost window is an “XY-plot” window created by using either the “Scatter plot” command or the “Regression line” command, this menu item is changed to either “Hide curve fit” or “Show curve fit” depending on the state of the window. If the plot already contains a curve fit, you can use the “Hide curve fit” command to hide this curve fit. On the other hand, if the plot does not contain a curve fit, then you can use the “Show curve fit” command to show it.

Show/hide statistics

To the right of the main plot area in an “XY-plot”, DynRisk displays the basic statistics for the corresponding pair of data nodes: “Slope”, “Intercept”, “Covariance”, and “Correlation”.

If the frontmost window is an “XY-plot” window, this menu item is changed to either “Hide statistics” or “Show statistics” depending on the state of the window. If the plot already contains a legend, you can use the “Hide legend” command to hide this. Conversely, if the plot does not contain a legend, then you can use the “Show legend” command to show it.

Note that this menu item is also used for editing “S-curves”, “Inverse S-curves” and “Histograms” and is then called “Hide/Show legend”.

Curve fit settings…

DynRisk calculates the curve fit by modeling the relation between the two data nodes as a so-called Gaussian process. You can control how this is done by adjusting parameters such as “Resolution”, “Smoothness” and “Sensitivity”. To do this you select the node you want to adjust the parameter settings for and then use the “Curve fit settings…” command.

Resolution

This parameter determines the density of fitted points along the curve.

Smoothness

By selecting a high smoothness factor, the fitting algorithm will try to find a “smoother” curve, i.e., local “jumps” will be straightened out. When using a low smoothness factor the algorithm interprets every “jump” as significant, and the curve will typically be very unstable.

Sensitivity

This parameter determines how sensitive the algorithm is to large deviations from the mean value trend.

Note! Usually it is a good idea to select a rather high smoothness factor (e.g., 90). If you have many data points, you may choose “High” resolution, otherwise choose “Medium” or “Low”. The sensitivity factor should be greater than 50, unless you have very many data points. Make sure that the scale of the X-axis is chosen such that there are data points all along the axis.