



Advanced Modelling with AgenaRisk

www.agenarisk.com

Copyright Agena Limited - © All rights reserved.

Copyright Notice

© 2005 Agena Ltd. No part of this publication may be reproduced, published, stored in an electronic database, or transmitted, in any form or by any means, electronic, mechanical, recording or otherwise, for any purpose, without the prior written permission of Agena Ltd.

THERE IS NO WARRANT OF ANY KIND FOR THE ACCURACY OR USEFULNESS OF THIS INFORMATION EXCEPT AS REQUIRED BY APPLICABLE LAW OR EXPRESSLY AGREED IN WRITING.

1	INTRODUCTION	4
2	MODELS WITH MULTIPLE RISK OBJECTS	5
2.1	Introducing the River Flooding Model	5
2.2	Importing Models	5
2.3	Connecting Risk Objects	7
2.4	Running the Model	8
3	ADVANCED USE OF NPT EXPRESSIONS AND SIMULATION	10
3.1	Introducing the Hypothesis Testing Model	10
3.2	Using Conditional Statements	10
3.3	Using Discrete Probability Distributions	13
3.4	Running the Model	15
3.5	Capturing Different Prior Assumptions	16
3.6	Combining Data and Prior Assumptions	18
3.7	Changing the Simulation Settings	19
4	NEXT STEPS	21

1 Introduction

This tutorial illustrates how AgenaRisk can be used to build advanced and complex models. In the first half of the tutorial, you will learn how to build and use models that consist of more than one connected risk object. In the second half of the tutorial, you will explore NPT expressions and simulation in depth by building a powerful hypothesis-testing model.

Before working through this tutorial you should ensure that you have a sound working knowledge of AgenaRisk by reading the previous two tutorials: *Getting Started with AgenaRisk* and *Modelling with AgenaRisk*.

Throughout this tutorial, various formatting conventions are used:



File	A term appearing in bold will sometimes refer to a menu on the toolbar that you are required to select.
OK	Bold is also used to indicate any button that you are required to click.
23	Bold can be used to indicate any value that you are asked to enter or select.
Menu → Menu Item	Two or more bold terms separated by an arrow indicate a path from a toolbar menu (or from the Start menu) down to a menu item that you are required to click.
Ctrl + W	The use of Ctrl, Alt or Shift in conjunction with a + symbol and a letter indicates that the key should be pressed down simultaneously with the letter to invoke a particular function.
<i>Filename.ast</i>	Names of files and directories (used during opening, saving and importing) are indicated by italics.
<i>Directory / Filename.ast</i>	Two or more italicised terms separated by a forward slash indicate a path from a directory (or folder) down to a particular file.
	This icon indicates a tip or other piece of useful information.
	This icon indicates a warning; you should pay particular attention to these.

Table 1 Formatting conventions

2 Models with Multiple Risk Objects

2.1 Introducing the River Flooding Model

1. The River Flooding model predicts whether a river will flood using information about the water level of a river, the amount of rainfall and the flood defences in place to prevent flooding.
2. The model is structured in such a way that is possible to link together different instances of it to monitor the flood risk over time.
3. The amount of rainfall, together with the “prior” water level (i.e. before rainfall), influences the “post” water level (i.e. after rainfall). This “post” water level can then be used as the “prior” water level in a new instance of the model that represents the next time period.
4. The quality of flood defences and the water level after rainfall determine whether or not the river will flood. As with the “post” water level, the quality of the flood defences can be linked to the same node in the next time period.

2.2 Importing Models

1. Start AgenaRisk and open the *Examples / Tutorials / River Flooding Basic Object.ast* model.
2. Ensure that the *River Flooding* risk object is selected in the Risk Explorer. The model should look like the one in Figure 1 below.

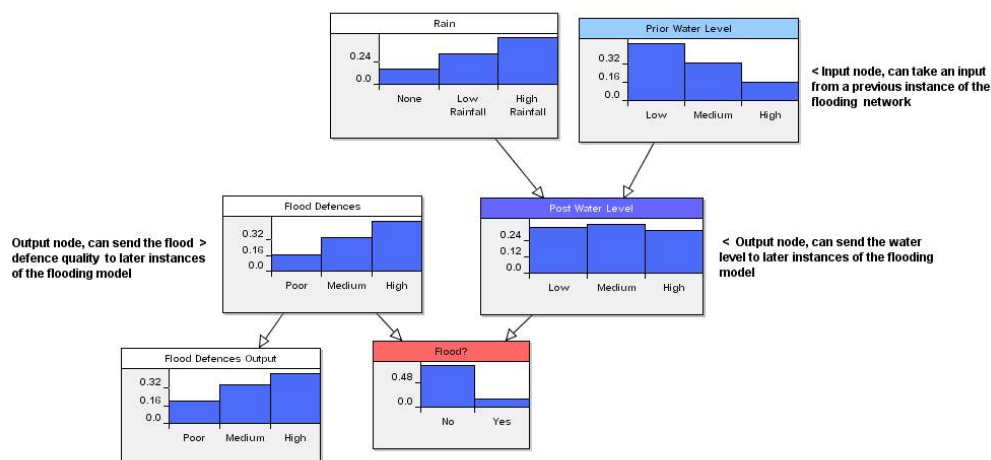


Figure 1 River Flooding model

3. Experiment with different values of *Rain*, *Prior Water Level* and *Flood Defences* to see what impact they have on the probability of a flood occurring.
4. When you are comfortable with how the basic model works, open the properties dialog for *Prior Water Level*, check the *Input Node* box and then click **OK**.
5. Repeat step 4 for *Flood Defences*.
6. Open the properties dialog for *Post Water Level*, check the *Output Node* box and click **OK**.

7. Repeat step 6 for *Flood Defences Output*.
8. Now save the modified model and call it *Dynamic River Flooding.cmp*.
9. Click on **File** → **Import Model...**, select the *Dynamic River Flooding.cmp* file that you just saved and click **OK**.
10. Click **No** in the dialog that asks you about applying the properties of the imported model to its nodes.
11. The Risk Explorer now appears and you can see both models:

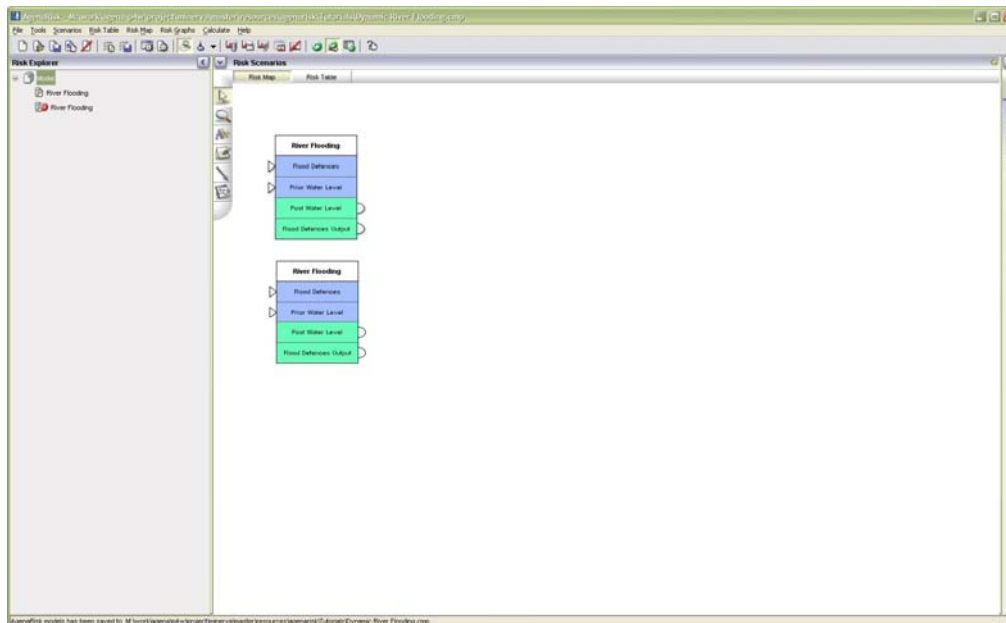


Figure 2 *River Flooding* model showing two risk objects after import

12. Notice that the Risk Map shows both risk objects and that each risk object lists its input nodes (shown in blue) and output nodes (shown in green).
13. Import *River Flooding.ast* once more so that your model contains three identical risk objects.
14. In the Risk Explorer, rename the three objects *Day 1*, *Day 2* and *Day 3*. Then, in the Risk Map, line up the three objects horizontally so that they appear in chronological order. Your model should now look like the one shown in Figure 3.

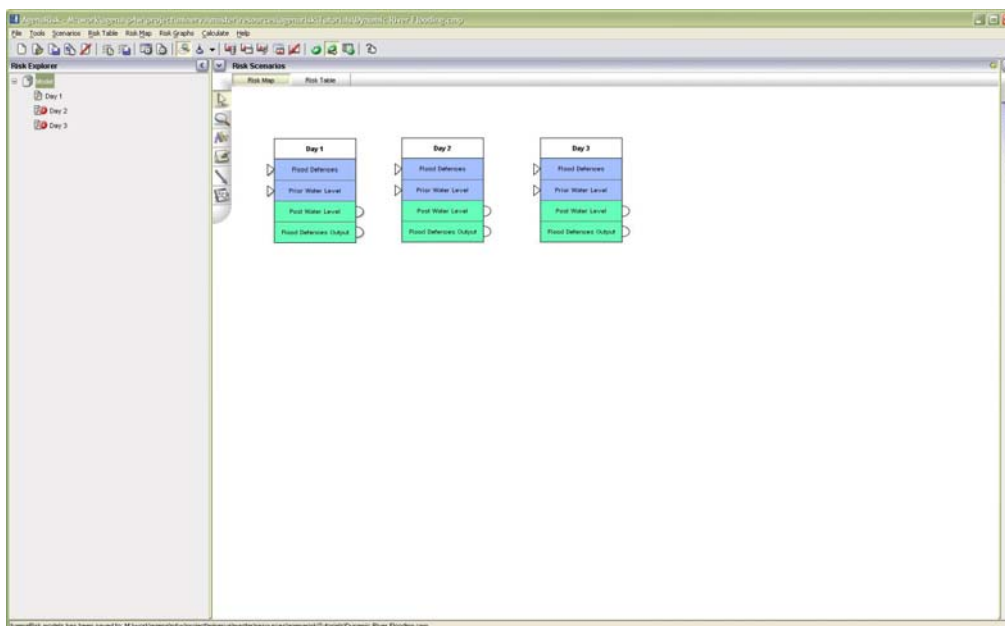


Figure 3 River Flooding model showing three risk objects

15. The River Flooding model now contains three risk objects that are identical in all but name. Select each risk object in the Risk Explorer and verify that the Risk Map and Risk Table for each object look the same.

2.3 Connecting Risk Objects

1. In order to capture the time dimension in our model so that we can reason about how changes in the water level each day affect the risk of flooding, we need to connect the three risk objects together.

2. Click on the edge tool on the Risk Map toolbar:



3. Click on the *Post Water Level* segment of the *Day 1* risk object, move your mouse to the *Prior Water Level* segment of the *Day 2* risk object and click again. This creates a connection between the two risk objects:

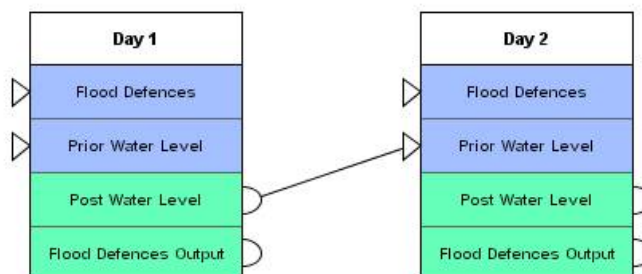


Figure 4 A link between two risk objects

4. This link means that, whenever the model is calculated, the probabilities associated with each state of the *Post Water Level* node in *Day 1* are entered as probability distributions in the *Prior Water Level* node in *Day 2*. In this way, the water level at the end of the first day becomes the water level at the beginning of the second day.

5. Draw another link between *Flood Defences Output* in *Day 1* and *Flood Defences* in *Day 2* so that the two risk objects look like those in Figure 5.

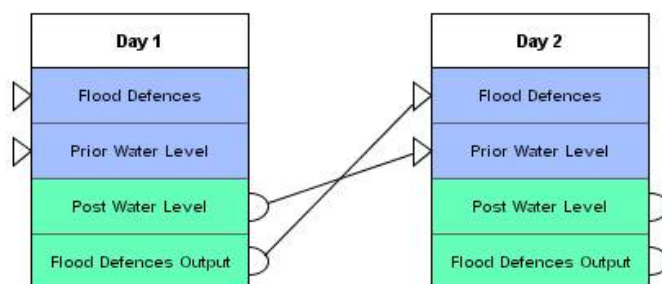


Figure 5 Two connections between two risk objects

6. Connect *Day 2* with *Day 3* in the same way:

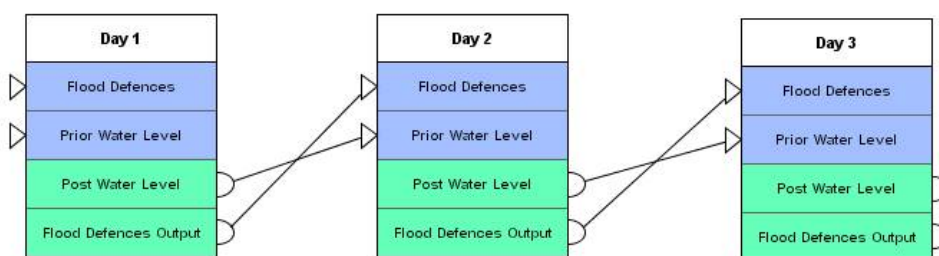


Figure 6 Three connected risk objects

7. The model is now fully connected and can be used to assess how the risk of flooding changes over a three-day period.

2.4 Running the Model

1. Turn on the auto-calculate option by toggling the toolbar button:



2. Close all risk graphs by selecting **Risk Graphs** → **Close All Graphs** from the menu bar or by clicking:



3. You are now going to use the model to explore how three days of high rainfall affects the risk of flooding and the quality of the flood defences.
4. Using the Risk Explorer, click on **Day 1** and enter the observation **High Rainfall** on the *Rain* node.
5. Enter the same observation for **Day 2** and **Day 3**.
6. Return to **Day 1** in the Risk Explorer and double-click on the *Flood?* and *Flood Defences* nodes to show their risk graphs. Hold your mouse over the bars on the graphs to view the exact probability values of the states.

7. Click on **Day 2** in the Risk Explorer and show the same two risk graphs. Notice that the risk of flooding has increased and that the quality of the flood defences has degraded.
8. Click on **Day 3** in the Risk Explorer. Open the two risk graphs and observe that the risk of flooding has increased again and that the quality of the flood defences has degraded further still.
9. Suppose, now, that you want to see what effects remedial action on day 3 will have on the risk of flooding.
10. Ensuring that **Day 3** is still selected in the Risk Explorer, enter the observation **High** on the *Flood Defences* node.
11. Verify that the risk of flooding on day 3 has now been reduced by over **half**.
12. You might like to check your model against the solution which can be found in *Examples / Tutorials / River Flooding.ast model*

3 Advanced Use of NPT Expressions and Simulation

3.1 Introducing the Hypothesis Testing Model

1. Imagine we have two materials, A and B. From sampled data we wish to find which of the two has the best quality; that is, which material has the lowest probability of containing faults.
2. Let's say that we take 200 samples of material A and 100 samples of material B. Of these we find that 10 of material A and 9 of material B are faulty.
3. Based on this data, can we truly believe that material A is better than material B? We need to test the so-called 'null hypothesis':

$H_0: p_A = p_B$ (meaning there is no difference in the quality of the materials)

against the so-called 'alternative hypothesis':

$H_1: p_A > p_B$ (meaning material A is better than material B)

3.2 Using Conditional Statements

1. Create a new model and open up the model properties dialog by clicking on **File** → **Model Properties**.
2. Check the boxes marked **Use a Continuous X-Axis** and **Treat Min/Max X as percentiles**.
3. In the **Lower percentile X Bound**, enter the value **0.1** and in the **Upper percentile X Bound**, enter the value **99.9**.
4. Your model properties dialog should look like the one shown in Figure 7 below.

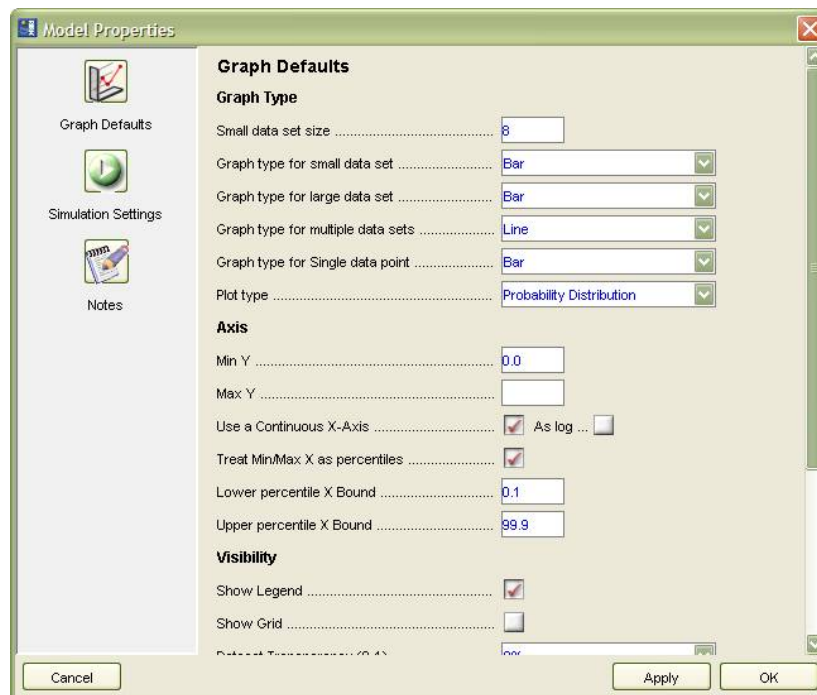


Figure 7 Properties for *Hypothesis Testing* model

5. Click **OK**.
6. These steps ensure that, by default, risk graphs for any nodes that we create will be continuous and that the extreme tails of distributions will not be plotted.
7. Expand the Risk Explorer so that you see the blank Risk Map for the single risk object in the model. (You can rename the risk object in the Risk Explorer if you like.)
8. In this model, we are going to use the following colour scheme for nodes:

Light Blue	Sample Nodes
Medium Blue	Probability Nodes
Dark Blue	Result Node
Green	Assumption Node

Table 2 Colour scheme for *Hypothesis Testing* model

9. Add a new node and change the following properties:

Node Name	Probability Material A is Faulty
Unique Identifier	p_material_a_faulty
Simulation Node	Ticked
States	Single state: 0 – 1
Background Colour	Medium Blue

Table 3 Properties for new node

10. Click **Apply** and then click **OK** on the information message that appears (telling us that a Uniform expression has been created automatically).
11. Switch to the **Node Probability Table** tab and verify that **Uniform** is selected in the drop-down box, that the **Lower Bound** field contains the value **0.0** and that the **Upper Bound** field contains the value **1.0**.
12. Click **OK**.
13. Copy the node you just created, paste the copy and move it so that it appears to the right of the original.
14. Change the **Node Name** of the new node to **Probability Material B is Faulty** and change the **Unique Identifier** to **p_material_b_faulty**.
15. Copy either of the two nodes and paste the copy beneath them.
16. Make the new node a child of the two existing nodes.
17. Change the **Node Name** of the new node to **p(A is faulty) – p(B is faulty)** and change the **Unique Identifier** to **a_b_difference**.
18. Make the lower bound of the single state **-1.0** instead of **0.0**.

19. Change the type of expression on the node from **Uniform** to **Arithmetic** and, in the **Arithmetic Expression** text box, supply the following expression:

p_material_a_faulty – p_material_b_faulty



While you are typing the expression, you will notice that a red border surrounds the box. This indicates that the expression is invalid. If you hold your mouse over the box, a tooltip will appear that explains the cause of the problem. When the expression is complete and valid, the red border will disappear.

20. Create a new node below $p(A \text{ is faulty}) - p(B \text{ is faulty})$. This is going to be a conditional node that represents our two hypotheses and shows the probability of each of them being true.
21. Make the new node a child of $p(A \text{ is faulty}) - p(B \text{ is faulty})$.
22. Open the properties dialog for the new node.
23. Change the **Node Name** to **Hypothesis**, change the **Unique Identifier** to **hypothesis** and change the **Node Type** to **Boolean**.
24. Click **Apply**.
25. In the **Node States** tab, type **Material A better than Material B** in the **Positive Outcome** field and type **Material A not better than Material B** in the **Negative Outcome** field. These two states correspond to the alternative and null hypotheses we want to test.
26. Click **Apply**.
27. Switch to the **Node Probability Table** tab and change the **Node Editing Type** to **Expression**. Because this is a Boolean node, only one type of expression is available: Comparative.
28. In the **Comparative Expression** field, supply the following expression:

**if(a_b_difference<0,
"Material A better than Material B",
"Material A not better than Material B")**

This expression means that, during calculation, the **Material A better than Material B** state will be chosen when the difference between the two probabilities is less than zero; otherwise, the **Material A not better than Material B** state will be chosen.



Make sure that you type the exact, case-sensitive values of the states. Otherwise, when you generate the NPT, you will get an error message telling that you the NPT contains zero probabilities (because the NPT generator has been unable to match the states in the expression with the states of the node).

29. Click **Apply**.
30. In the **Appearance** tab change the background colour of the node to dark blue and in the **Text Format** tab, change the text colour to white.
31. Click **Apply**.
32. In the **Graph Defaults** tab, click the inheritance symbol



next to the following fields: **Use a Continuous X-Axis, Treat Min/Max X as percentiles, X-Axis start state** and **X-Axis end state**.

33. Uncheck the **Use a Continuous X-Axis** and **Treat Min/Max X as percentiles** boxes and clear the **X-Axis start state** and **X-Axis end state** text boxes.
34. Click **OK**.
35. Run the model, display the risk graph for the *Hypothesis* node and verify that, in the absence of any information, the probability of material A being better than material B is slightly below **0.5**. Your model should now look similar to the one in Figure 8 below.

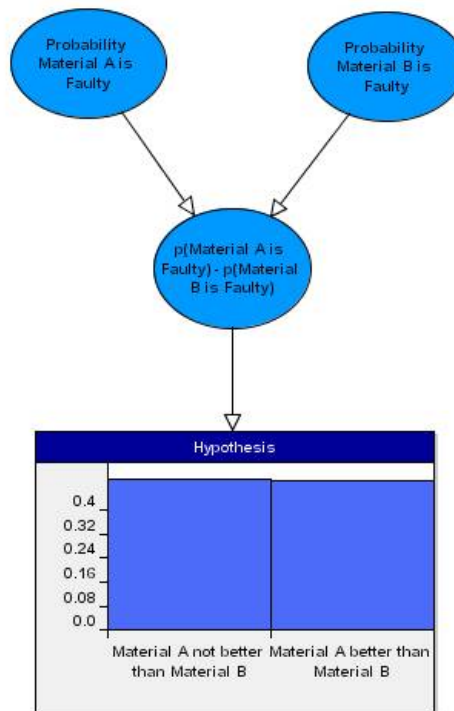


Figure 8 Probabilities of *Hypothesis* node with no observations

3.3 Using Discrete Probability Distributions

1. You are now going to add nodes that represent the samples of the two materials taken during the test.
2. Create a new node and make it a child of *Probability Material A is Faulty*.
3. Open the properties dialog for this node.
4. Change the **Node Name** to **Material A: faults in 200 trials** and change the **Unique Identifier** to **a_faults**.
5. Make the node a simulation node and change its type to **Integer Interval**. This type is appropriate because the number of faults observed during sampling is always going to be a whole number.
6. Click **Apply**.

- Switch to the **Node States** tab, click on the **Remove all states** hyperlink and then check the **Make Upper Bound Positive Infinity** box. You should now have two states as shown in Figure 9.

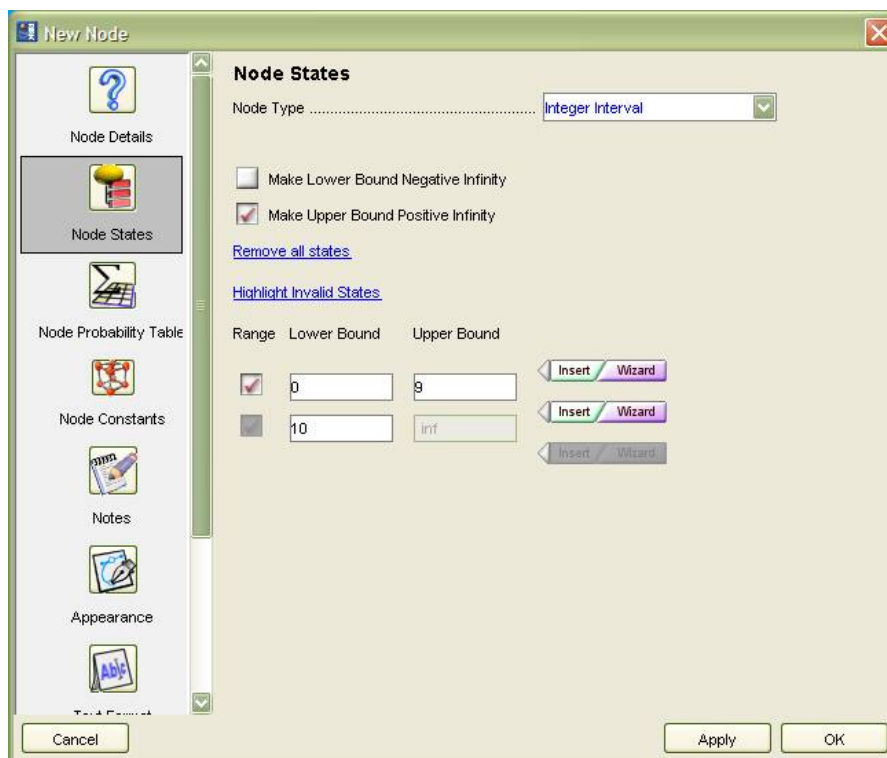


Figure 9 States on a new Integer Interval node

- Click **Apply**.
- Click on the **Node Probability Table** tab, change the **NPT Editing Mode** to **Expression** and select **Binomial** in the drop-down box.
- Enter the value **200** in the **Number of Trials** text box and enter the unique identifier of this node's parent (**p_material_a_faulty**) in the **Probability of Success** text box. This means that the NPT generated for this node will be based on a Binomial distribution of faults observed. When you come to use the model, you will enter hard evidence on this node (in the form of a number of faulty samples observed). Then, via backward calculation, the model will update the probability of material A being faulty based on this evidence.
- Click **Apply**.
- In the **Appearance** tab change the background colour of the node to light blue.
- Click **OK**.
- Copy *Material A: faults in 200 Trials* and paste it. When you copy and paste a node that has one or more parents, the new copy is connected to the same set of parents by default.
- Select the edge between *Probability Material A is Faulty* and the new node and delete it.
- Move the new node over to the right of the Risk Map and make it a child of *Probability Material B is Faulty*.

17. Bring up the properties dialog for the new node.
18. Change the value of **Node Name** to **Material B: faults in 100 trials** and change the value of **Unique Identifier** to **b_faults**.
19. Click on the **Node Probability Table** tab. Notice that there is a red warning saying that the NPT needs to be regenerated and that there is a red border around the **Probability of Success** text box. This is because the Binomial expression refers to the parent of the copied node. It needs to be rectified. Clear the **Probability of Success** box and enter **p_material_b_faulty**.
20. In the **Number of Trials** box, enter the value **100**.
21. Click **OK**.
22. The model is now complete and should look similar to the one in Figure 10.

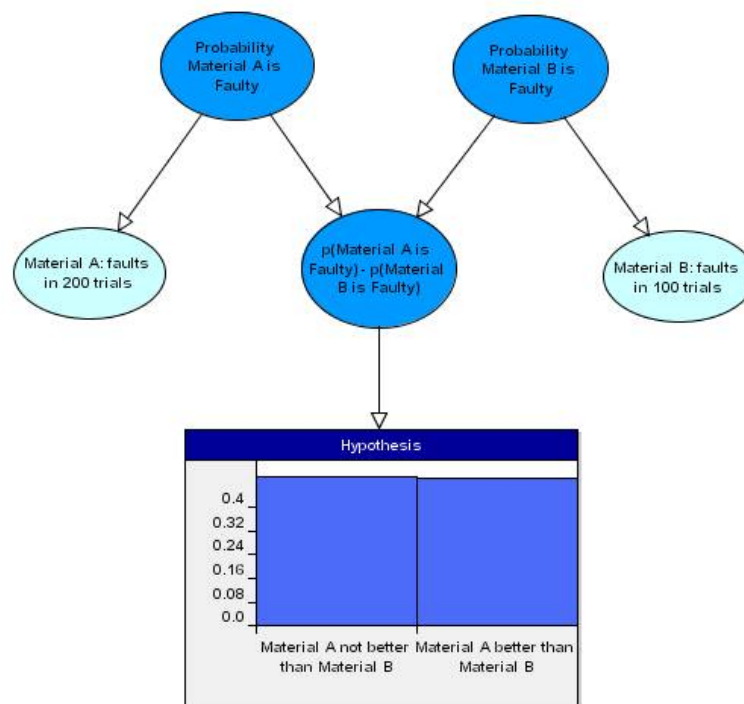


Figure 10 Complete *Hypothesis Testing* model

3.4 Running the Model

1. You can now use the model to test the two hypotheses.
2. If automatic calculation is not already activated, turn it on by toggling the toolbar button:



3. Enter the observation **10** for the node *Material A: faults in 200 trials* and enter the observation **9** for *Material B: faults in 100 trials*.
4. Look at the risk graph for the *Hypothesis* node. It should resemble the one in Figure 11 below.

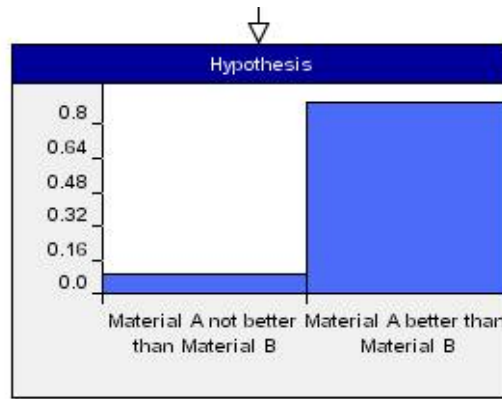


Figure 11 Risk graph of *Hypothesis* after evidence has been entered

5. Hold your mouse pointer over the bar on the right and verify that the probability of material A being better than material B is **0.90437** (i.e. roughly 91%).
6. Experiment with different numbers of samples to see how the result of the hypothesis test changes.
7. Save your model.
8. At this point you might want to compare your model against the solution which can be found in *Examples / Tutorials / Hypothesis Testing.ast*

3.5 Capturing Different Prior Assumptions

1. Close down any open risk graphs.
2. Create a new node at the top of the Risk Map. (Select all of the existing nodes and move them down if necessary.)
3. Open the properties dialog and change the following properties:

Node Name	Prior Type
Unique Identifier	prior_type
Node Type	Labelled
States	Uniform, Beta
Background Colour	Green

Table 4 Properties for new node

4. Click **Apply**.
5. In the **Graph Defaults** tab, click the inheritance symbol



next to the following fields: **Use a Continuous X-Axis**, **Treat Min/Max X as percentiles**, **X-Axis start state** and **X-Axis end state**.

6. Uncheck the **Use a Continuous X-Axis** and **Treat Min/Max X as percentiles** boxes and clear the **X-Axis start state** and **X-Axis end state** text boxes.

7. Click **OK**.
8. Make *Prior Type* a parent of both *Probability Material A is Faulty* and *Probability Material B is Faulty*.
9. Open the properties dialog for *Probability Material A is Faulty* and switch to the **Node Probability Table** tab.
10. Change the **NPT Editing Mode** to **Partitioned Expression** and move the **Prior Type** entry in the left-hand list over to the right-hand list. You are going to supply one expression to be used when the state of *Prior Type* is *Uniform* and a different expression to be used when the state of *Prior Type* is *Beta*.
11. Double-click on the first cell in the expression table and, in the dialog that appears, specify a **Uniform** distribution with a **Lower Bound** of **0** and an **Upper Bound** of **1**. This is equivalent to a prior assumption of complete ignorance about the likelihood of material A being faulty.
12. Click **OK**.
13. Double-click on the second cell in the expression table and, in the dialog that appears, specify a **Beta** distribution. Enter **1** for **Alpha** and **9** for **Beta**. Specify a **Lower Bound** of **0** and an **Upper Bound** of **1**. This is equivalent to a prior assumption that there is a 1 in 10 chance of material A being faulty.
14. Click **OK** and click **OK** again to close the properties dialog.
15. Define a partitioned expression on *Probability Material B is Faulty* just as you did for *Probability Material A is Faulty*. Specify it identically but make the **Alpha** parameter of the **Beta** distribution **2** and the **Beta** parameter **8**. This encodes the assumption that there is a 1 in 5 chance of material B being faulty.
16. Clear the two existing observations from the model (by choosing **Tools** → **Clear Entered Data** → **All** from the menu bar).
17. Double-click on the two nodes whose NPTs you just modified and resize the nodes so that you can see the shapes of the graphs. Notice how the two prior assumptions differ:

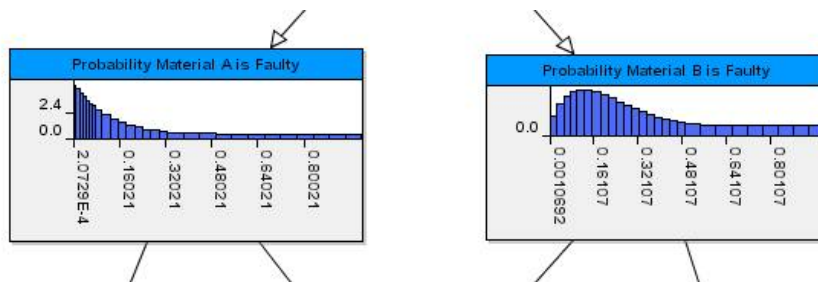


Figure 12 Hypothesis Testing model showing different prior assumptions

18. Close the two risk graphs.
19. You are now going to compare two scenarios side-by-side: one in which the prior assumption is that the distribution of faults is uniform and the other in which a beta distribution of faults is assumed. The latter scenario might represent the opinion of a materials expert, for example.

20. Expand the Risk Scenarios panel and rename **Scenario 1** to **Uniform**.
21. Add a new scenario, rename it **Beta** and tick the **Active** checkbox.
22. Right-click on the *Prior Type* node and choose **Enter Observation** → **Uniform** → **Uniform**. By doing this, you are entering the observation that the *Prior Type* is *Uniform* in the scenario that you have named **Uniform**.
23. Right-click on the *Prior Type* node again and choose **Enter Observation** → **Beta** → **Beta**. By doing this, you are entering the observation that the *Prior Type* is *Beta* in the scenario that you have named **Beta**.
24. Double-click on the *Hypothesis* node to show its risk graph. It should look like the one shown in Figure 13.

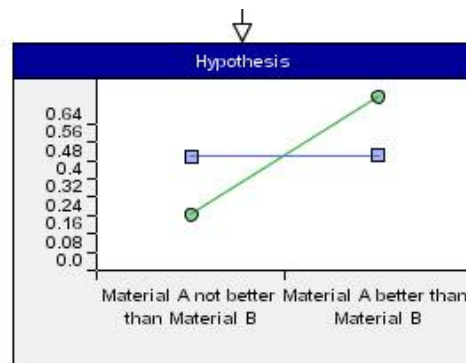


Figure 13 Results of hypothesis test with two different prior assumptions

Note that the two scenarios are plotted on it. Drag your mouse the points on the two lines and note the values of the two states for each scenario.

25. The data tells us that, in the absence of any sampling information, when we assume that the probabilities of faultiness are uniformly distributed, there is roughly a **50%** chance that material A is better than material B.
26. However, if our materials expert believes that the prior faultiness probabilities are characterised by beta distributions – Beta(1,9) and Beta(2,8) for material A and material B respectively – then this results in a **75%** chance that material A is better than material B.

3.6 Combining Data and Prior Assumptions

1. Open up the properties dialog for the node *Material A: faults in 200 trials*.
2. Change the name of the node to *Material A: faults in 10 trials* and change the **Number of Trials** parameter of the Binomial distribution from **200** to **10**.
3. Click **OK**.
4. Open up the properties dialog for the node *Material B: faults in 100 trials*.
5. Change the name of the node to *Material B: faults in 10 trials* and change the **Number of Trials** parameter of the Binomial distribution from **100** to **10**.
6. Click **OK**.
7. Turn off the auto-calculate feature (because you are about to enter four observations as a “batch” and you don’t need to calculate the model after each one).

8. For both the **Uniform** and the **Beta** scenario, enter an observation of **1** fault on *Material A: faults in 10 trials*.
9. For both the **Uniform** and the **Beta** scenario, enter an observation of **0** faults on *Material B: faults in 10 trials*.
10. Run a calculation. The graph for the *Hypothesis* node should now look like this:

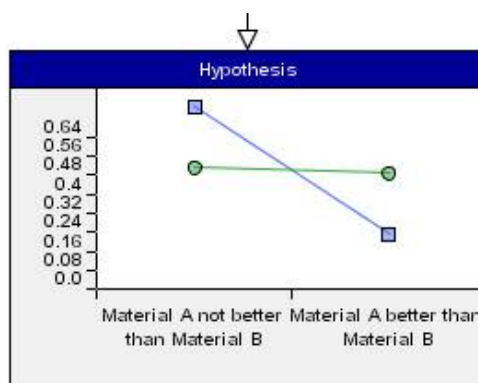


Figure 14 Results of hypothesis test after entering sparse sample data

11. Hold your mouse over the points of the graph to see the state probabilities.
12. Notice that, when we use uniform priors for the two faultiness probabilities, the sparse sample data tells us, with a confidence of roughly **77%**, that material A is not better than material B. However, when we encode the opinion of an expert in the two Beta priors, these priors serve to supplement the sample data and tell us that the quality of the two materials is roughly **equal**.
13. Save your model.
14. At this point you might want to compare your model against the solution which can be found in *Examples / Tutorials / Hypothesis Testing with Expert Judgement.ast*

3.7 Changing the Simulation Settings

1. Close all the risk graphs.
2. Delete the **Uniform** scenario (by right-clicking on it in the Risk Scenarios panel and choosing **Delete Selected Scenario(s)**).
3. Remove the observation of **1** fault from *Material A: faults in 10 trials*. You can do this by right-clicking on the node and choosing **Enter Observation** → **Beta** → **Clear Observation**.
4. Run a calculation.
5. Open the risk graph for *Material A: faults in 10 trials*. Drag your mouse over the graph and observe that the mean number of faults is **1.0078** and that the variance is **1.6932**.
6. Open the model properties by clicking on **File** → **Model Properties** from the menu bar and switch to the **Simulation Settings** tab.
7. Verify that the current value in the **Maximum Number of Iterations** field is **25**. This setting determines how long you want simulation to continue for. The longer simulation goes on for, the more accurate it will be.

8. Change the value of **Maximum Number of Iterations** to **5** and then click **OK**.
9. Run a calculation.
10. Observe that the calculation takes less time but the results are less accurate: the mean number of faults is now **1.6437** and the variance is **1.2277**.
11. Modelling often requires you to make a trade-off between speed and accuracy; in some situations accuracy will be more important and in others speed will be more important.

4 Next Steps

This tutorial has shown you some of the advanced aspects of building models using AgenaRisk. To learn more about the many types of problem AgenaRisk can be used to model, you should explore and interact with the range of well-documented example models that come with the software.