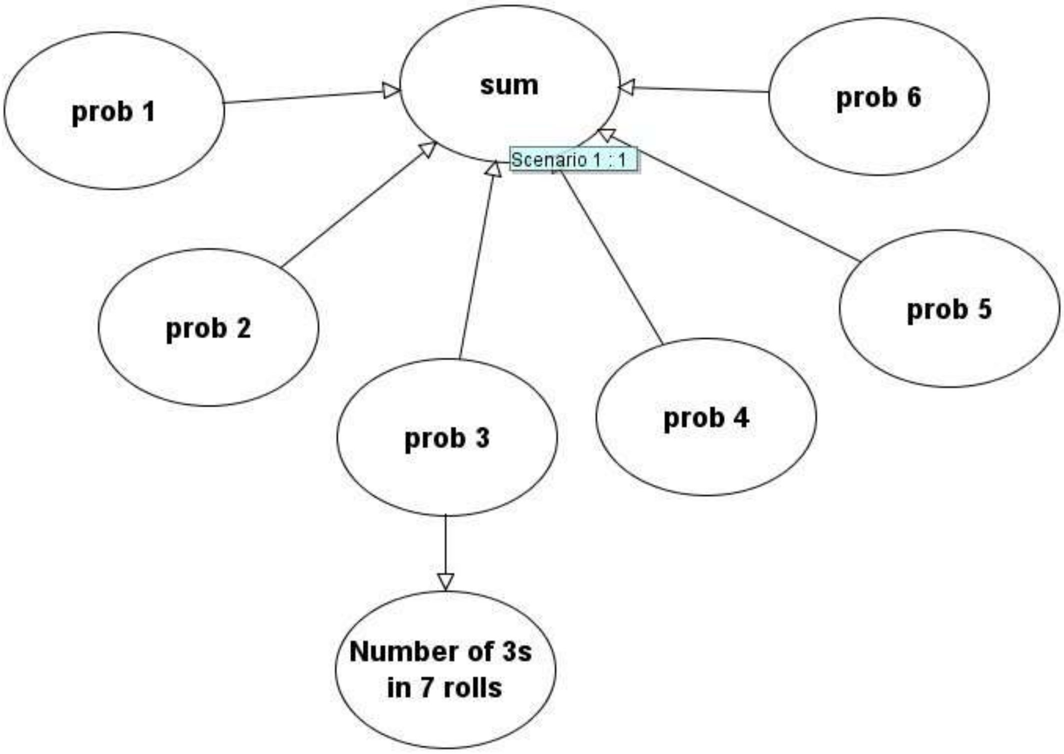
What is the chance the next roll of the die will be a three

Consider the following question:

> The die has rolled 3 3 3 3 3 3 3 in the past. What are the chances of 1 2 4 5 6 being rolled next? The mathematician will say: P(k)=1/6 for each number, forget that short-term evidence. What will the probability expert say? And the statistician? And the philosopher?

We can use a Bayesian solution to provide the answer in each case (with the possible exception of the 'statistician' curiously enough). The model we need (it is a Bayesian network) is this one:



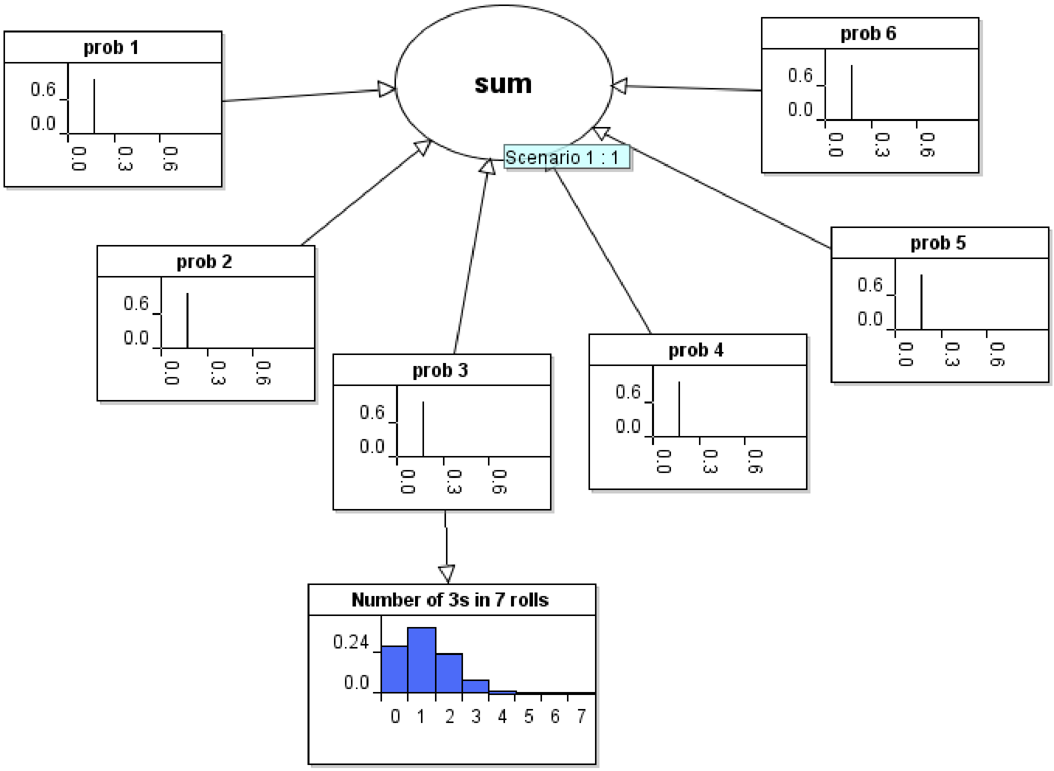
Here the nodes prob 1, prob 2, ..., prob 6 represent the probability of rolling the number 1,2,...,6 respectively.

The node 'sum' is a logical constraint on the model (it is the sum of the six probabilities and, because of the probability axioms this sum must be 1 assuming that no outcome other than 1,2,...,6 is possible from rolling the die).

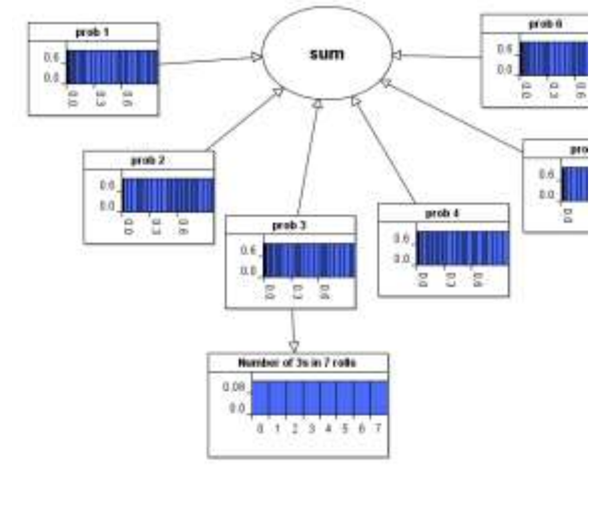
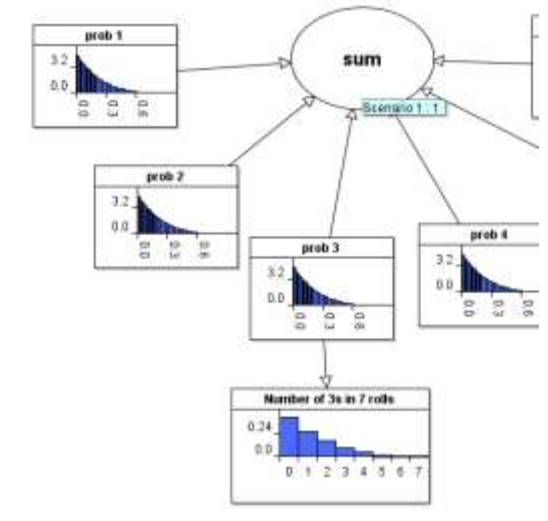
The node "number of 3's in 7 rolls" is defined as a Binomial distribution (the probability that the number is 7 is simply $p3^7$ where p3 is the probability of getting a 3).

In the model we have to set some prior probability distribution on each of the nodes prob 1, prob 2, ..., prob 6 (the particular choice of prior is what distinguishes the mathematician, probability expert and philosopher as well as layman). Before we enter the evidence of the 7 rolls of 3, the model - when calculated - displays the prior marginal probabilities. Thus:

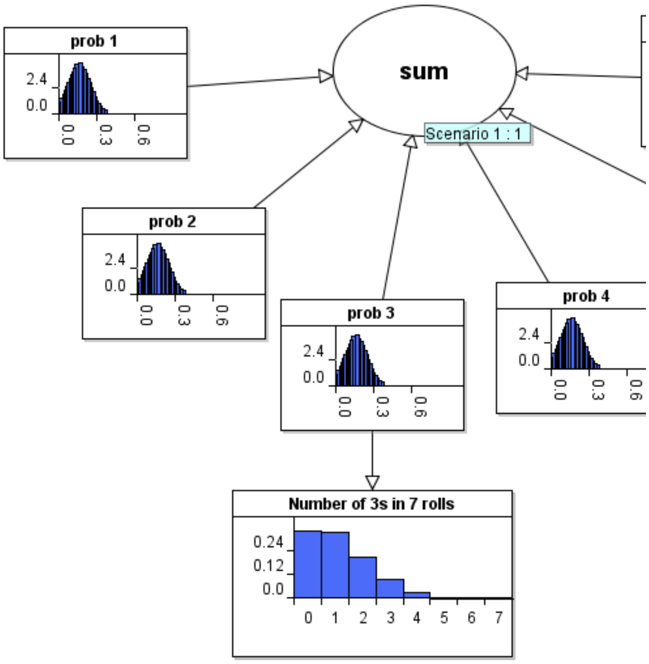
The mathematician's prior is that the probability of each Pk is ***exactly*** 1/6, so the prior probability distributions Pk looks like this:



One type of probability expert (including certain types of Bayesians) will argue that, in the absence of any prior knowledge of the die, the probability distribution for each Pk is ***uniform*** over the interval 0-1 (meaning any value is just as likely as any other), so the prior probability distribution for each node Pk looks like this (before and after we enter the sum=1 constraint):

	
Before sum constraint	after sum constraint

Another probability expert (including most Bayesians) will argue that the prior should be based on dice they have previously seen. They believe most dice are essentially 'fair' but there could be biases due to either imperfections or deliberate tampering. Such an expert might therefore specify the prior distribution for Pk to be a narrow bell curve[1] centred on 1/6:

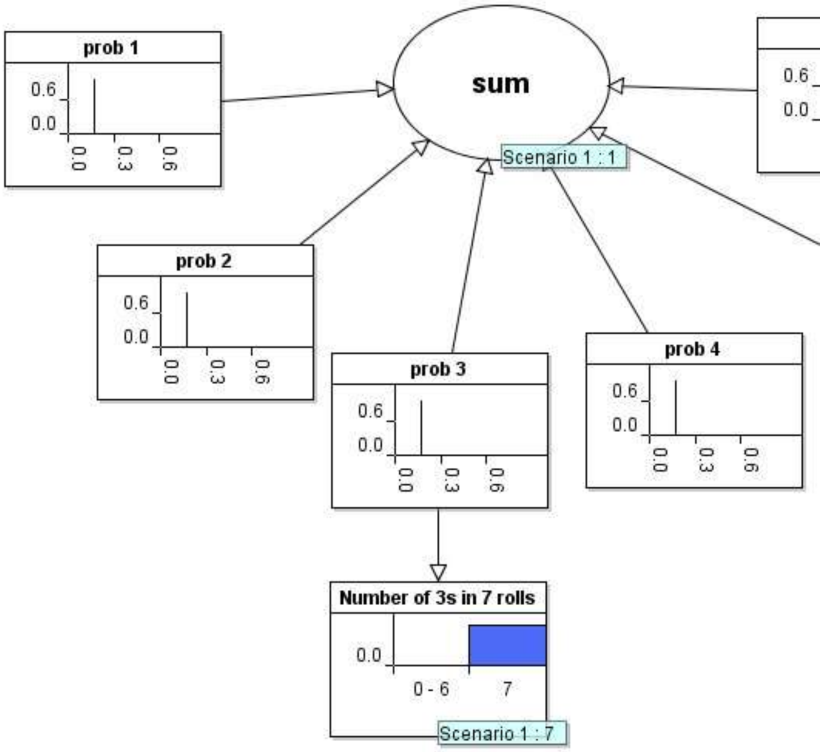


A philosopher might consider any of the above but might also reject the notion that 1,2,3,4,5,6 are the only outcomes possible.

Anyway, when we enter the evidence of seven 3's in 7 rolls, the Bayesian calculations (performed here using AgenaRisk) result in an updated posterior distribution for each of the nodes prob 1, ..., prob 6:

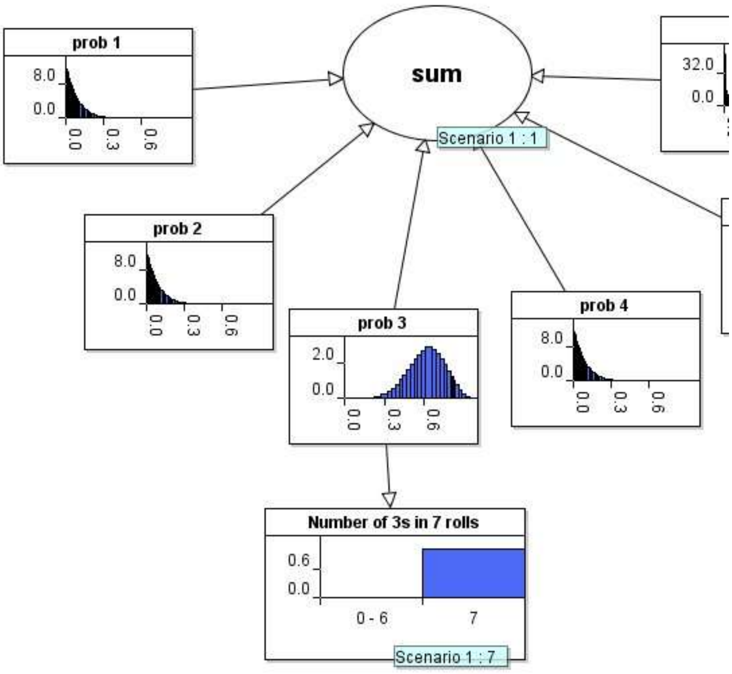
[1] We have used a Truncated Normal distribution with mean 1/6 and variance 0.01 over the range 0-1

The mathematician's posterior for each node is unchanged:



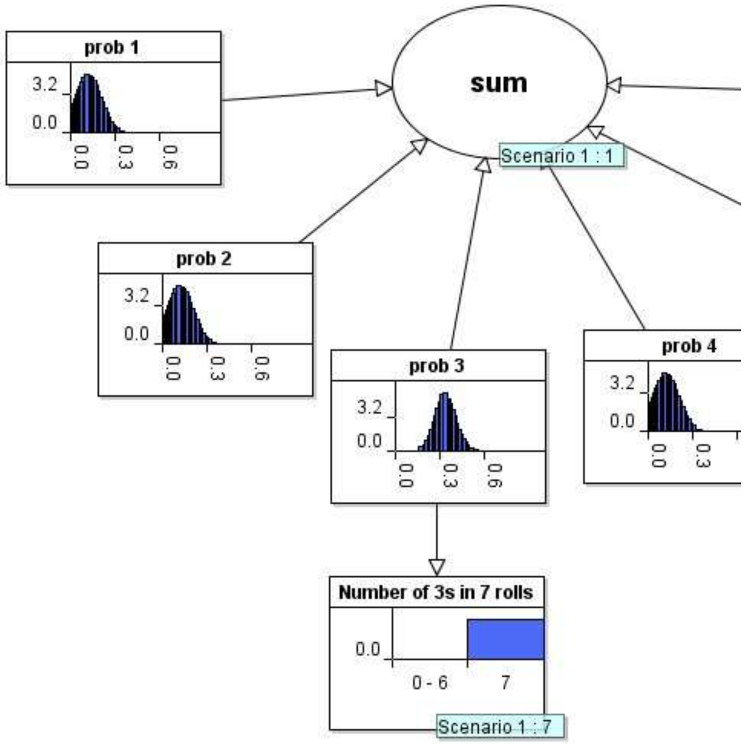
i.e. the probability the next roll of the die will be 1,2,3,4,5,6 are all respectively still 1/6.

The posteriors for the probability expert with the uniform priors:



The prob 3 is now a distribution with mean 0.618. The other probs are all reduced accordingly to distributions with mean about 0.079. So in this case the probability of rolling a 3 next time is about 0.618 whereas each of the other numbers has a probability about 0.079

The posteriors for the probability expert with bell curve priors:



The prob 3 is now a distribution with mean 0.33. The other probs are all reduced accordingly to distributions with mean about 0.13. So in this case the probability of rolling a 3 next time is about 0.33 whereas each of the other numbers each has a probability about 0.13.

And what about the statistician? Well a classical statistician cannot give any prior distributions so the above approach does not work for him. What he might do is propose a 'null' hypothesis that the die is 'fair' and use the observed data to accept or reject this hypothesis at some arbitrary 'p-value' (he would reject the null hypothesis in this case at the standard p=0.01 value). But that does not provide much help in answering the question. He could try a straight frequency approach in which case the probability of a three is 1 (since we observed 7 out of 7 threes) and the probability of any other number is 0.