University of Illinois at Urbana-Champaign

Research Projects Irisk Lab

Visualization Of Sample Recycling Methods for Nested Stochastic Modeling

Prepared by:
Hao Gai
Jeremy Soriano
Samuel Woessner
Wilson Phurwo
Zhangyao Chen

Mentors:
Haoen Cui
Runhuan Feng

December 5, 2018

Abstract

As the need of accuracy and efficiency develops in stochastic modeling, the new Sample Recycling Method is used to save more time by recycling a limited set of scenarios within a certain error range. Since this technique is commonly used in lots of fields not only in economics but also in engineering, telecommunication and so on. For more and more people having a deeper understanding, we created a video in a vivid way to introduce Sample Recycling Method for Nested Stochastic Modeling.

1 Introduction

Explanation of the sample recycling method to general audience has been a serious challenge. The complexity of the mathematics behind it requires simplified interpretation in order to understand the underlying concept. There are many techniques that can be done to approach this concern, but it heavily depends on the targeted audience. The most convenient assumption of this audience is to be as generic as possible. In addition, it is reasonable to assume that they have, at least, slight background in mathematics. Therefore, we chose to approach this problem by visualization by animated video, which hopefully delivers digestible content. For more advanced purpose, our team also produces an interactive web page that was developed by Samuel Woessner.

The rest of the paper is organized as follows: Section 2 provides a brief introduction of how we produce the video; Section 3 mainly talks about the Nested Monte Carlo Simulations and Sample Recycling Method by one example about the harvesting crops profits differently in sunny and rainy days; Section 4 concludes the paper and give some more applications in other fields.

2 Methodologies

As actuarial science students, we approached sample recycling method by understanding the concept at the first place. The research paper written by Haoen Cui, Runhuan Feng, and Peng Li guides us through this process. Once the concept is understood, we came up with a simple example that would be included in the video. The purpose of this example is to allow the audience of the video catch the big picture of sample recycling method within a short period of time, i.e. the length of the video. We began writing a script on Google Drive so that all of us can edit it simultaneously. The script went through a rigorous checking because we had to make sure that every fact stated in the script is true and valid. This process involved our team and Professor Runhuan Feng. The script was converted to storyboard which helped us pre-visualize the video content. Once the script was finalized, we used an online animation making platform, powtoon, to create the visualization. Besides the animated video, our interactive web page provides a more technical example, designed especially for people with actuarial science background or that of any related field. This example covers explanation of how nested Monte Carlo simulation and sample recycling method are used to evaluate stock price. This webpage was developed in R.

3 Simulations

3.1 Background

With all the uncertainty in our day to day lives, would not it be nice to know what to expect in the future? While we cannot completely predict the future, we can look at all possibilities and try to prepare for different scenarios. Using the computational tools we have today, we are able to estimate future stock prices, understand the dynamics of disease pathways to help treat illnesses, or even to estimate your saving toward retirement when considering different economic scenarios in the future.

However, Such computations require the use of running simulations, a process of computing hundreds, if not thousands, of scenarios to see what the most likely outcome of events are. An example of a simulation is the Monte Carlo simulation, and it is used in a large number of applications, such as climate studies, physics, engineering, finance, and even law. A simple example of a Monte Carlo simulation is throwing two dice. Suppose you want to know how likely it is to get a seven as a result of throwing two dice, meaning if you get one and six, two and five, and so on, you will count it as a success. Rather than using complicated mathematics, we can estimate this probability by looking at this chart, the probability of getting a seven is simply $\frac{1}{6}$.

However, Monte Carlo simulation requires us to actually throw two dice multiple times. For instance, you throw the dice one hundred times, and observe how many times seven comes up. You will see that the result will be close to $\frac{1}{6}$. If you do the simulation one thousand times then the result will eventually converge to $\frac{1}{6}$.

3.2 Nested Monte Carlo Simulations

Now that we understand how a Monte Carlo simulation works, we can begin to explain the idea behind a nested Monte Carlo Simulation. The nested simulation is different from an ordinary simulation in that we consider multiple layers including the inner loop and outer loop. We often estimate quantities of interest with a range of outer loop scenarios over a period of time, during which additional quantities require further estimations from inner loop simulations. Let us explain these ideas with an example.

Suppose we are trying to find the expected payoff from harvesting crops, based on the probability of the weather for the next 2 days. As a farmer, we are expecting to harvest in 2 days and can only harvest crops when it is not raining outside. If the weather is sunny we can harvest and make a profit of \$1 per crop, if it is raining, we are not able to harvest and we make a profit of \$0.

Using historical weather data, we are able to find that the future weather is dependent on the current weather. We also know that it is currently sunny today. With this information we can create a weather model for the next two days to calculate the expected payoff from harvesting.

If it is sunny today, there is a 60% chance that it will also be sunny tomorrow and a 40% chance that it will rain. The same probabilities would apply for day 2 as well, if it is sunny on day 1. The probability of weather in next day is denoted by P_s if it is sunny today.

$$P_s = \begin{cases} 0.6, & \text{sunny next day} \\ 0.4, & \text{rainy next day.} \end{cases}$$

However, if it rains on day 1, the chance of it being sunny tomorrow is 30% and the chance of it raining again the next day is 70%. The probability of weather in next day is denoted by P_r if it is rainy today.

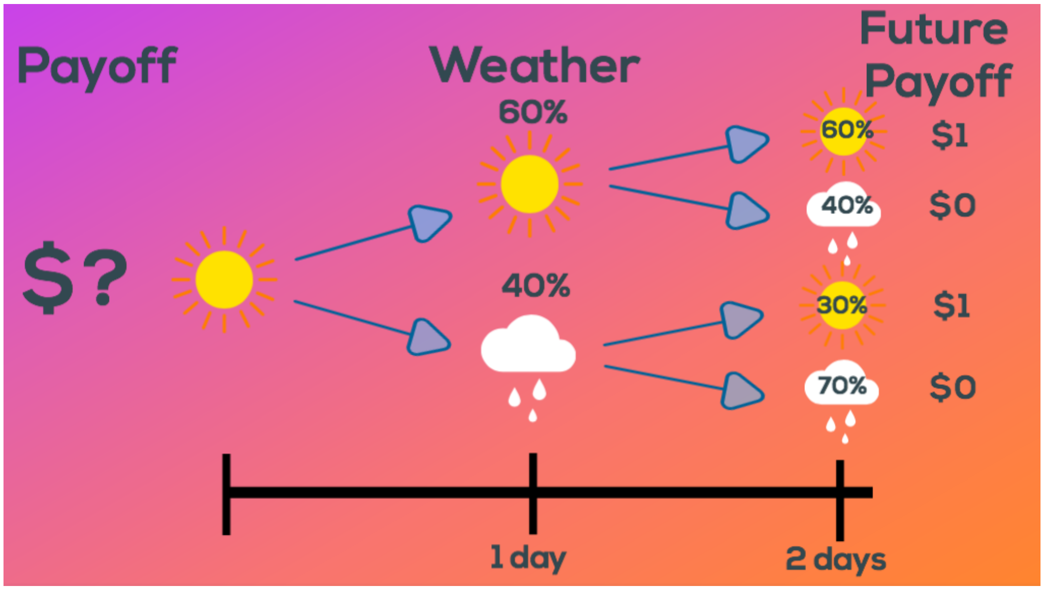

Figure 1: Expected Profits by Nested Monte Carlo Simulation

$$P_r = \begin{cases} 0.3, & \text{sunny next day} \\ 0.7, & \text{rainy next day.} \end{cases}$$

In order to calculate the payoff we would expect to receive today, we must take into consideration all possibilities the weather can take for the next two days and assume the risk free rate is 0. We can write the payoffs we would receive for each outcome as well, \$1 if sunny at the end of 2 days and 0 otherwise. Using these probabilities, we can now run a simulation to get a sample of outcomes. We must run a simulation for both paths in our model, let us start with the case of it being sunny tomorrow. The simulation runs a trial based off the probabilities in our model, giving us an idea of what we are expected to receive. In this case we run 10 simulations to generate weather outcomes, Resulting in 6 sunny days with a payoff of 1 on each day and 4 rainy days with a payoff of 0, corresponding with the probabilities with our model. With this sample, we calculate the expected payoff we would receive today by taking the average of all the payoffs in our simulation, resulting in an expected payoff of .6 or 60 cents per crop by $\frac{1}{10}(1 \times 6 + 0 \times 4)$.

We would do the same for the other path as well. Running another simulation based off the probabilities of the other path; Running a simulation with these probabilities, we get a sample of 3 sunny days with a payoff of 1 and 7 rainy days with a payoff of 0, taking the average once again we get an expected payoff of 30 cents per crop for this path by $\frac{1}{10}(1 \times 3 + 0 \times 7)$.

The term to describe these paths represented by a sample of simulations are sample paths. In order to calculate the expected payoff today, we multiply the expected payoffs we calculated, with the respective probabilities of each path occurring, resulting in an expected payoff of 48 cents per crop at time 0.

$$\mathbb{E}[\text{Payoff}] = 0.6 \times 0.6 + 0.3 \times 0.4 = 0.48.$$

In this example we only ran the simulation 10 times to get our sample of 10, but in practice we must run the simulation to get a much larger sample size in order to be accurate.

Note further that we only had 2 sample paths representing sunny and rainy weather. In practice we would need many sample paths representing a multitude of economic scenarios and run a large simulation for them as well, increasing the amount of calculations exponentially for each new sample path we have.

In order to help reduce the amount of calculations needed, we can use the Sample Recycling Method, which uses only one sample path to calculate the expected outcome of all the possible alternative paths.

3.3 Sample Recycling Method

To explain how Sample Recycling works, let us use the same example as before. This time, using the same example as before, let us calculate the expected payoff the farmer would receive today using the sample recycling method.

In this method, we only use one sample from one path as a reference to calculate the expected outcomes of the other paths. Let us use the path that it will be sunny tomorrow as a reference to estimate the expected payoff of the path that it will rain tomorrow. We will refer to these paths as the reference and target paths. Next, let us find the expected payoff for each path, starting with the reference path.

Running through the same process as before, we obtain a sample of 10 outcomes for the scenario that it will be sunny tomorrow. We get a sample of 6 sunny days with a payoff of 1 and 4 rainy days with a payoff of 0. Hence we get an expected payoff of .6 for this path. Now we want to estimate the expected payoff for the target path.

This time, instead of running another simulation to get a new sample, like what we did in the Nested Monte Carlo Method, we reuse the sample of outcomes we had from the sunny scenario. In order to properly recycle this sample, we must find a relationship between the probabilities of the outcomes for the reference path and the target path.

We can do this by taking the ratio of the probabilities for each path and each outcome. Let us first find the ratio of probabilities for the outcome that it will be sunny. To do this, we take the target path probability of it being sunny (.3) and divide this by the reference path probability of it being sunny (.6) to get a ratio of .5, which we denote as λ_s.

$$\lambda_s = \frac{0.3}{0.6} = 0.5.$$

We do the same for the probability it will be rainy to get a ratio of 1.75, denoted as λ_r.

$$\lambda_r = \frac{0.7}{0.4} = 1.75.$$

These ratios are what we call the weight and we apply these to their respective outcomes in the reference sample. This adjusts the outcomes in our reference sample relative to the ratios between the reference and target path probabilities. $\lambda_s = .5$ means that the sunny outcome will occur half as many times in a hypothetical target sample than it will in our reference sample. Applying these ratios allows us to have a new sample that we can use for the target path to calculate its expected outcome without having to run another simulation.

Let us recall the reference sample (6 sunny samples and 4 rainy samples). Using the outcomes in our sample, we multiply the λ_s, the ratio of the target and reference probabilities that it will be sunny, with the payoffs we get if it is sunny. We do the same with the payoffs we get if it is rainy, using λ_r for those outcomes. Summing up all the outcomes and taking the average, we get an expected payoff of .3 for the target path.

$$\mathbb{E}[\text{ At Day 1 target path payoff}] = \frac{1}{10}(0.5 \times 6 \times 1 + 1.75 \times 4 \times 0) = 0.3.$$

Now that we have the two expected payoffs for each scenario, we can multiply their respective probabilities to get an expected payoff of .48, the same result as if we were to use the Nested Monte Carlo method.

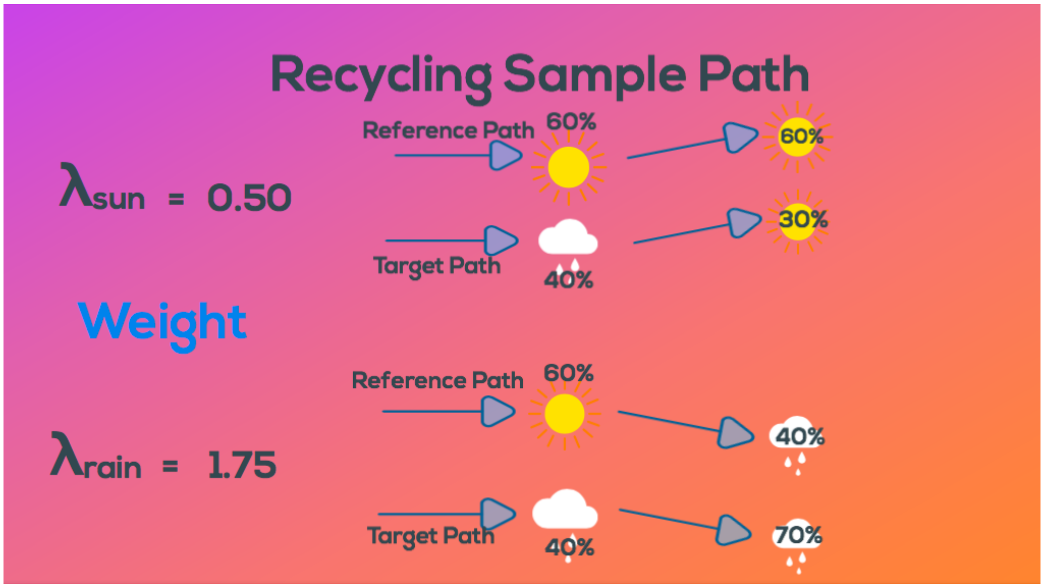

Figure 2: Sample Recycling Paths

As you can see, in this method we did not have to run a simulation to calculate another sample to find an expected payoff. We simply use the same sample and relate the probabilities of each outcome in the sample instead, using the lambdas we found. Again, in this example we only had 2 paths, but in practice we would need to have multiple to be accurate. Traditionally, with more scenarios requires more simulations to get a multitude of samples, requiring extensive calculations, but with the sample recycling method, we only require one sample, greatly reducing the amount of calculations needed.

4 Other Applications and Conclusions

As mentioned earlier, Monte Carlo simulation method is heavily used in many fields. In meteorology, weather forecasting often makes use of this technique to predict future atmospheric conditions. In telecommunication, wireless network design primarily depends on users? data, such as their location and the services they wish to use. Again, due to the randomness of users, Monte Carlo simulation helps telecommunication companies plan their design. In finance, stock prices are uncertain but one can estimate the cost of financial instruments whose prices depend on stock prices by running either a nested

Monte Carlo simulation or using the sample recycling method. To see this process in more detail, be sure to check our interactive web page.

To process uncertainty, many techniques can be done to draw a conclusion. Nested Monte Carlo simulation is accurate but requires a large number of simulations to get a result. To be more efficient, the sample recycling method is introduced and can be applied to everything that can be analyzed by nested Monte Carlo.

References

[1] Runhuan Feng. 2015 Society of Actuaries Survey on Nested Stochastic Modeling. 2015.

[2] Christian P. Robert and George Casella. *Monte Carlo Statistical Methods.* Springer, 2nd edition, 2004.

[3] Epstein, E.S. (December 1969). *"Stochastic dynamic prediction"*. Tellus A. 21 (6): 739-759. Bibcode:1969Tell...21..739E. doi:10.1111/j.2153-3490.1969.tb00483.x.

[4] Masashi Sugiyama, Taiji Suzuki, and Takafumi Kanamori. *Density Ratio Estimation in Machine Learning.* Cambridge University Press, 1st edition, 2012.

[5] Kroese, D. P.; Taimre, T.; Botev, Z. I. (2011). *Handbook of Monte Carlo Methods.* John Wiley & Sons.

[6] Peter G. Zhang. *Exotic Options: A Guide to Second Generation Options.* Singapore: World Scientic Publishing, 2nd edition, 1998.