Probabilistic Modeling

Nick DeRobertis

December 4, 2020

1 Introduction to Probabilistic Modeling

- Probability is a key concept for financial models as it related to risk

- The base result from a deterministic model only gives a single answer, but does not consider the probability distribution of the result

- E.g. your model could predict a positive NPV from a project, but through probabilistic modeling you determine that there is a 98% chance the NPV is negative. Do you still want to take the project? This is important information to know when making that decision.

- We discuss three techniques in this course that take advantage of probability theory: scenario modeling, internal randomness, and Monte Carlo simulation

- Scenario modeling and Monte Carlo simulation are also methods of exploring the parameter space, just like sensitivity analysis

- Internal randomness is useful for when the probability is so core to the model that it should be built in from the beginning, rather than extending the base model to add probability

2 Math Review for Probabilistic Modeling

- Discrete variables: specific values, continuous variables: range of values. Note that an underlying variable can be continuous but the modeler can choose to make it discrete within the model to simplify it. E.g. condition of the economy is continuous, but you might make it discrete by classifying the economic conditions into recession, neutral, or expansion. This does not work in the other direction: a variable which is truly discrete cannot be made continuous

- Expected value is a key concept in probability theory. We will use it in scenario analysis to get the expected outcome across multiple different cases of the inputs

- The variance graph shows two series with the same mean but different variances. Variance is a measure of how much the value is moving around. If it moves a lot, it has high variance. Variance has nothing to do with the average, mean, or expected value.

- Probability distributions tell you how likely it is to observe different values of a given variable

- Discrete variables: think table of values with probabilities. Continuous variables: think curve, usually displayed by a graph but defined by a function

- The CLT is extremely powerful, because of it most of the distributions for continuous variables are normal distributions. So if we need to pick a distribution for a variable, the normal distribution is reasonable choice the majority of the time

3   Introduction to Scenario Analysis

- Scenario modeling is another way to explore the parameter space, just like sensitivity analysis

- Scenario modeling is carried out in the same way as sensitivity analysis: run the model with each set of inputs and display the outputs with the inputs

- The difference is that unlike in sensitivity analysis where we tweak one input at a time, in scenario analysis, we consider some situations and determine all the values of the inputs that align with each situation

- We can also optionally assign probabilities to the situations, then take an expected value to get an expected outcome from our model

- Internal vs. external scenario analysis is just about whether the scenarios are included in the core model (internal), or as an extension to the base model (external)

4   Scenario Analysis in Excel

- We will focus on external scenario analysis for the purpose of this exercise

- Typically if the model has already been created, you should opt for external rather than internal

- Unfortunately without some hacks, in Excel we are limited to examining two inputs changing at once with external scenario analysis, as a data table is used to run the model repeatedly

- We will see that we do not have this limitation in Python, and later in the course we will see that you can use Python to run external scenario analysis on your Excel model with any number of inputs

5   Lab Exercise - Adding Scenario Analysis to Project 1 - Excel

- Here we are adding external scenario analysis to the Project 1 Excel model

- The process should be basically identical to what was shown in extending the Dynamic Salary Retirement Model in Excel

6   Scenario Analysis in Python

- Just as with sensitivity analysis and any other extension to the base model, your model needs to be set up appropriately to add scenario analysis. All the main logic should be in functions, and the functions should not access variables which were defined outside a function unless they are passed in as arguments.

- Dictionaries are a helpful data structure for scenario analysis as you can put the name of the scenario as the key and have the input values as the value

- The process is very similar to other methods of exploring the parameter space. Run the model with each set of inputs and keep the outputs associated with the inputs. In scenario analysis we often also assign probabilities to the cases and take the expected value of the results

- Before you complete the lab, ensure that your Python sensitivity analysis lab is working properly, or else this and any other future extensions to the model will probably also not work properly

7 Introduction to Internal Randomness

- Internal randomness can be very powerful but also has one major drawback: your model is no longer deterministic, so the same inputs will produce different outputs each time. While this is the goal of internal randomness, it also makes it difficult to know if the model is working correctly.

- You should only use internal randomness in cases where the randomness is so integral to the model that it wouldn't make sense to have the model without it

- The external counterpart to internal randomness is Monte Carlo simulation, which will be the next extension we examine for our models. In most cases, Monte Carlo simulation will achieve your goals and you don't need internal randomness

- Pick distributions which make sense for your inputs, but usually just pick a normal distribution with reasonable mean and standard deviation

- For the mean, use the historical mean or your best estimate of the expected value of the input. For the standard deviation, a rule of thumb is that you should pick a standard deviation such that +/- 1x stdev changes from the mean are common, 2x occur occasionally, 3x are rare, and 4x will almost never occur.

8 Intro to Randomness in Excel

- Excel has the functions to get most of the kinds of random numbers that you would want, but often multiple have to be combined such as NORM.INV(RAND()) for randomly drawing numbers from a normal distribution

9 Intro to Randomness in Python

- Python has far more and more convenient features around randomness than Excel

- A single function, random.normalvariate, can be used for generating random numbers from a normal distribution

- There is a lot better support for other conveniently using other distributions in Python

- You can also generate many random numbers at once using numpy

10 Lab Exercise - Generating Continuous Random Numbers in Excel and Python

- The exercise is the same for Excel and Python

- Make sure to do the entire exercise with n=10 and n=1000 so you can see the differences that number of iterations makes

11 Discrete Randomness

- The story around generating random discrete values in Excel is relatively complicated. In Python, it is just as straightforward as working with continuous numbers

- We can hack our way into a discrete random value function in Excel by using uniform random numbers and the probability distribution. It is a manual effort to set this up in your model

- In Python, just pass your discrete values and probabilities to the random.choices function and you are done. This function can even give you multiple choices

12   Adding Internal Randomness to an Excel Model

- Typically you would add internal randomness as you build the model, not afterwards. It can require substantial changes to the original logic to add it afterwards

- In most cases Monte Carlo simulation should be used when the model is already built

- Here we will be randomly picking for each year whether it is a recession, normal, or expansion economy

- In this case, as we want to add different behavior year by year based on the randomness, it is not possible via an external method such as Monte Carlo simulation

- Here I am treating interest rate and savings rate as discrete random variables since they are tied to the state of the economy which is a discrete random variable. I could have assigned the quality of the economy to a normal distribution to make it continuous and then the other variables could be continuous as well

- There are no special tricks here in Excel, we are just applying the previous approach for discrete random variables and hooking it into the model

13   Adding Internal Randomness to a Python Model

- See the notes for Adding Internal Randomness to an Excel Model as the conclusions are the same here

- It is quite an effort to add internal randomness which is this integral to the model after it has already been built

- Similarly to in Excel, there are not any new concepts in this lecture, but we are just applying the discrete random variables concepts and integrating into the model

14   Internal Randomness Lab Exercises Overview

- Whereas I was adding internal discrete randomness in the examples, now you will add internal continuous randomness to your own model

- I have selected the investment return as the random variable as it is only used in the NPV calculation. It should only require minimal changes to your model to integrate this

- This is a case where we would normally use Monte Carlo simulation as we are just randomizing the existing inputs. Here we are just applying internal randomness to understand how to implement it