# Monte Carlo Simulation

## Analyzing Probabilities of Outputs

Nick DeRobertis[1]

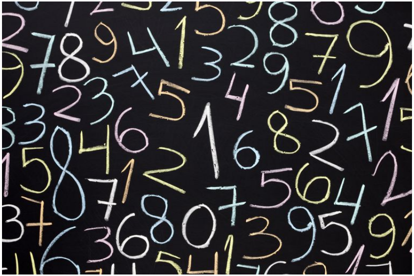
[1]University of Florida
Department of Finance, Insurance, and Real Estate

December 4, 2020

# Table of Contents

# What is Monte Carlo Simulation?

- **Monte Carlo Simulation** is a technique which allows understanding the probability of acheiving certain outputs from a model.

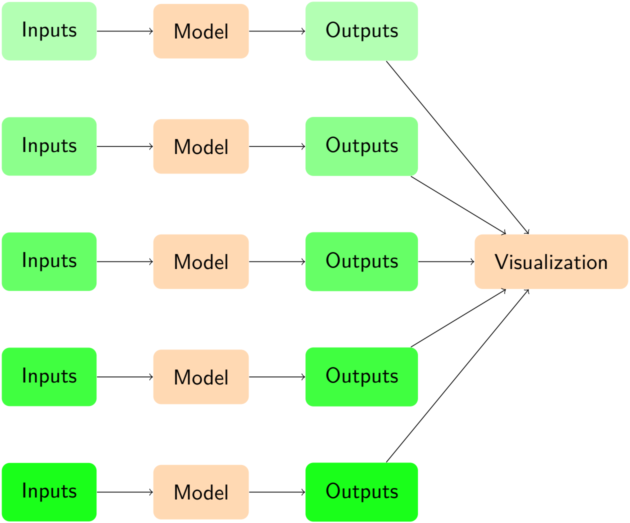


- This gives the modeler a greater understanding of the likelihood of different outputs, rather than relying on a single number

# Why Use Monte Carlo Simulation?

- Imagine you have a one-time opportunity to place a bet for \$1.
- If you win the bet, you will receive \$2. If you lose the bet, you will lose \$750,000. You cannot avoid the payment by declaring bankruptcy.
- The odds of winning the bet are 999,999/1,000,000. In 1/1,000,000 you lose the \$750,000.
- The expected profit from the bet is \$0.25. Should you take it? Depends on your risk tolerance.
- Therefore not only the expected outcome matters, but also what other outcomes may occur and their probabilities.

# Monte Carlo Simulation in One Picture



# Basic Process for Monte Carlo Simulation

- Monte Carlo simulation is carried out similarly to external scenario analysis.
- The main difference is that we manually picked specific cases for the inputs with scenario analysis.
- In Monte Carlo simulation, we assign distributions to the inputs, and input values are drawn from the distributions for each run of the model
- Finally, we can fit a probability distribution to the outputs to be able to talk about the chance of a certain outcome occurring

# Table of Contents

# Running Monte Carlo Simulations - Python or Excel?

- Monte Carlo simulation can be applied to any model
- It is generally easier to run them in Python than in Excel.
- With pure Excel, you're either going to VBA or hacking something with data tables
- In Python, just loop for N iterations, each time drawing inputs, running the model, and collecting outputs.
- We will start with a pure Python model, then move to using `xlwings` to add Monte Carlo simulations to our Excel models.

# An Example Application

## An Investment Problem

You have \$1,000 now and need to pay \$1,050 in one year. You have available to you two assets: a risk free asset that returns 3%, and a stock that returns 10% with a 20% standard deviation. How much should you invest in the two assets to maximize your probability of having at least \$1,050 in one year?

- We must first construct the basic model which gets the portfolio value for given returns
- Then draw values of the stock return from a normal distribution, and run the model many times and visualize the outputs.
- Then repeat this process with each weight to determine the best weight.

# Simluating Portfolio Values

## Example for Simulating Portfolio Values

- Go to the course site and download the Jupyter notebook "MC Investment Returns.ipynb" from Monte Carlo Examples
- I will go through this example notebook to solve the problem from the prior slide.

# Intro Monte Carlo Lab, Level 1

## Monte Carlo Simulation of DDM, Level 1

1. You are trying to determine the value of a mature company. The company has had stable dividend growth for a long time so you select the dividend discount model (DDM).
2. $P = \frac{d_1}{r_s - g}$
3. The next dividend will be \$1, and your baseline estimates of the cost of capital and growth are 9% and 4%, respectively
4. Write a function which is able to get the price based on values of the inputs
5. Then you are concerned about mis-estimation of the inputs and how it could affect the price. So then assume that the growth rate has a mean of 4% but a standard deviation of 1%
6. Visualize and summarize the resulting probability distribution of the price

---

Level 2: Slide 30 Resources: Slide 31

# Table of Contents

# Monte Carlo Simulation Process

For the model given by:

$$y = f(X) \tag{1}$$

$$X = [x_1, x_2, ..., x_n] \tag{2}$$

- $y$ : Model output
- $X$ : Model input matrix
- $x_i$ : Value of $i$th $x$ variable

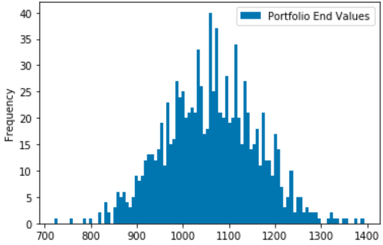
To run $N$ Monte Carlo simulations, follow the following steps:

1. Assign a probability distribution for each $x_i$
2. For each $x_i$ randomly pick a value from its probability distribution. Store them as $X_j$
3. Repeat the previous step $N$ times, yielding $[X_1, X_2, ..., X_N]$
4. For each $X_j$ calculate $y_j = f(X_j)$
5. Store the values of $X_j$ mapped to $y_j$
6. Visualize and analyze $y_j$ versus $X_j$

# Outputs from Monte Carlo Simulation

- There are a multitude of outputs we can get from a Monte Carlo simulation. We saw a few already in the example.
- **Outcome probability distributions** are the main output. We saw this with two approaches in the example, a histogram and a probability table.
- We also examined the **probability of a certain outcome** in whether we reached the desired cash.
- The last main output is examining the **relationship between inputs and outputs.** for which common approaches include scatter plots and regressions.

# Outcome Probability Distributions



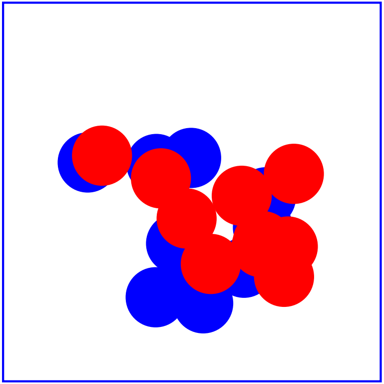

Probability Table

| Probability | Value |
|---|---|
| 25% | 1020 |
| 50% | 1039 |
| 75% | 1053 |

- The outcome probability distribution represents the chance of receiving different outcomes from your model.
- There are two main ways to visualize a probability distribution: a plot and a table.
- The plot, usually a histogram or KDE gives a high-level overview of the probabilities and can uncover any non-normal features of the distribution.
- The probability table represents the chance of receiving the given value or lower.
- The Value at Risk (VaR) is a common measure calculated in the industry, and it represents the probability of losing at least a certain amount. This would be a subset of this analysis and so this analysis can be used to calculate VaR

# Probability of a Certain Outcome - A Simple Example

- Imagine a box which contains red and blue balls. You do not know in advance how many there are of each color.
- You want to estimate the probability of getting a blue ball when pulling a ball from the box.
- To evaluate this, you grab a ball, write down its color, and put it back, 1,000 times.
- You pull a blue ball in 350 out of the 1,000 trials. What is the probability of getting blue?



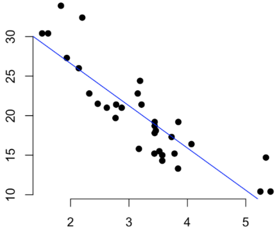
# Probability of a Certain Outcome, Formally

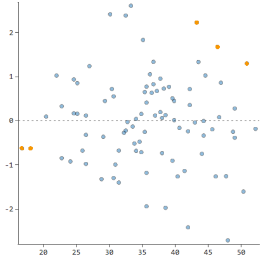
- We followed the same logic when estimating the probability of receiving our desired cash in the investment example.
- $p = \frac{\text{Count of positive outcomes}}{\text{Count of trials}}$
- For the balls example, this is simply $p = \frac{350}{1000} = 0.35$
- In the investment example, we used `pandas` to check for each trial, whether it was a positive outcome (made it a 1) or not (made it a 0). Then the sum is the count of positive outcomes and so the mean is the probability.

# Why Monte Carlo Simulations Help Understand Inputs vs. Outputs

- Monte Carlo simulation can also provide a more comprehensive look at the relationship between inputs and outputs.
- While sensitivity analysis can be used to estimate the relationship between an input and output, it is usually done with other inputs at their base case
- The values of inputs may affect how other inputs affect the output. E.g. for the retirement model, an increase in interest rate increases wealth more if the initial salary was higher.
- As all the inputs change each time, you can get a more realistic view of the relationship, e.g. some trials with a higher interest rate will have high salary and some will have low salary.

# Visualizing the Relationship between Inputs and Outputs




- A scatter plot is a simple way to visualize the relationship between two variables
- If there is a relationship, you will see some defined pattern in the points. This may be somewhat of an upward or downward line (linear relationship) or some other shape such as a U (non-linear relationship).
- If there is no relationship, then there will just be a random cloud of points (lower plot) or a horizontal line.

# Numerically Analyzing the Relationships

- The scatter plots help give a broad understanding of the relationship but do not answer the question, how much will my output change if my input changes? E.g. if I earn 10,000 more for a starting salary, how much sooner can I retire?
- Simply increasing the input in your model and checking the output is not enough, because it does not take into account how all the other variables may be changing.
- Multivariate regression is a general tool which is good at answering these kinds of questions, while taking into account all the changing inputs.

| | I | J | K | L | M | N | O |
|---|---|---|---|---|---|---|---|
| 5 | Regression Analysis | | | | | | |
| 6 | | | | | | | |
| 7 | OVERALL FIT | | | | | | |
| 8 | Multiple R | 0.919501 | | | | | |
| 9 | R Square | 0.845482 | | | | | |
| 10 | Adjusted R Square | 0.806853 | | | | | |
| 11 | Standard Error | 0.36759 | | | | | |
| 12 | Observations | 11 | | | | | |
| 13 | | | | | | | |
| 14 | ANOVA | | | | Alpha | 0.05 | |
| 15 | | *df* | *SS* | *MS* | *F* | *p-value* | *sig* |
| 16 | Regression | 2 | 5.914849 | 2.957425 | 21.88703 | 0.00057 | yes |
| 17 | Residual | 8 | 1.080978 | 0.135122 | | | |
| 18 | Total | 10 | 6.995827 | | | | |
| 19 | | | | | | | |
| 20 | | *coeff* | *std err* | *t stat* | *p-value* | *lower* | *upper* |
| 21 | Intercept | 0.17685 | 0.434522 | 0.406999 | 0.694681 | -0.82516 | 1.178858 |
| 22 | Color | 0.299435 | 0.051206 | 5.847594 | 0.000384 | 0.181352 | 0.417517 |
| 23 | Quality | 0.230012 | 0.047229 | 4.870187 | 0.00124 | 0.121103 | 0.338921 |

# How to use Multivariate Regression

- The coefficient in a multivariate regression represents how much the outcome variable changes with a one unit change in the input variable.
- E.g. a coefficient of -.0002 on starting salary in explaining years to retirement would mean that a $1 increase in starting salary is associated with a decrease in years to retirement by .0002 years, or a $10,000 increase in starting salary is associated with a decrease in years to retirement by 2 years.
- All interpretations are "all else constant", meaning that it does not consider relationships between the inputs. E.g. if starting salary is higher because of a good economy, and interest rates are also higher due to the good economy, the starting salary coefficient is not taking into account the increase in interest rates.
- Be careful about units. If you use decimals for percentages, you will need to multiply or divide by 100 to get the effect in percentages.

# Adding Monte Carlo Simulation to a Formal Model

## Dynamic Salary Retirement with Monte Carlo

- I will now go through adding a Monte Carlo simulation to the Dynamic Salary Retirement Model in Python
- The completed example is on the course site in Monte Carlo Examples

# Monte Carlo Python Lab, Level 1

## Monte Carlo Simulation of Python Models, Level 1

1. Work off of your existing Project 1 Python model
2. You are concerned the NPV could be heavily affected by changes in the interest rate. Instead of fixing it, draw it from a normal distribution with mean of 7% and standard deviation of 2%.
3. Run the model 10,000 times and collect the NPV results. Visualize the results. Create a table of probabilities and the minimum NPV we could expect with that probability. Output the chance that the NPV will be more than $400,000,000.

---

Level 2: Slide 32 Resources: Slide 33

# Table of Contents

# How is it Different Running MC in Excel?

- In pure Excel, it is much more difficult to run a Monte Carlo Simulation
- Without going to VBA, typically the only way is to use a data table
- A data table can be used in situations where you only want to have one or two inputs varying at once. Just generate the random inputs and use them as the axes of the data table
- If you want to vary more than two inputs, VBA or Python would be required
- There are also add-ons that accomplish this but they are usually not free

# Monte Carlo in Excel with More than Two Variables

- The process for Monte Carlo Simulation which works for any number of variables is very similar to what we were doing in Python.
- We are still just changing the inputs, running the model, and storing the outputs from each run
- Using `xlwings` from Python code we can change and retrieve the values of cells
- This allows us to change inputs, run the model, and store outputs, just as in Python, but running our Excel model.
- We can either analyze the outputs in Python or output them back to Excel for analysis

# Monte Carlo Excel Retirement Model

## Using `xlwings` to Run Monte Carlo Simulations

- Go to the course site and download the "Dynamic Salary Retirement Model.xlsx" and "Excel Monte Carlo.ipynb" from the Monte Carlo Examples
- Open up the Jupyter notebook and follow along with me
- The completed Excel model is also there in case you lose track. Visualizations were added after running the Jupyter notebook on the original Excel model.

# Monte Carlo Excel Lab, Level 1

## Monte Carlo Simulation of Excel Models, Level 1

1. You will be running Monte Carlo simulations on your existing Excel model from Project 1
2. You are concerned that your estimate for the number of phones that will be sold is incorrect.
3. The number of phones should instead be drawn from a normal distribution with mean 100,000 and standard deviation of 20,000.
4. Estimate the model 1,000 times and output the results back to Excel
5. In Excel, visualize the results. Create a table of probabilities and the minimum NPV we could expect with that probability. Output the chance that the NPV will be more than $400,000,000.

---

Level 2: Slide 34 Resources: Slide 35

# Lecture Resources

## Lecture Resources

1. Slides - Monte Carlo Simulation
2. Lecture Notes - Monte Carlo Simulation
3. Monte Carlo Investment Returns
4. Dynamic Salary Retirement Model with Monte Carlo
5. Dynamic Salary Retirement Model with Monte Carlo
6. Excel Monte Carlo

# Intro Monte Carlo Lab, Level 2

## Monte Carlo Simulation of DDM, Level 2

1. Continue from the first lab exercise
2. Now you are also concerned you have mis-estimated the cost of capital. So now use a mean of 9% and standard deviation of 2%, in addition to varying the growth
3. Visualize and summarize the resulting probability distribution of the price
4. Be careful as in some cases, the drawn cost of capital will be lower than the drawn growth rate, which breaks the DDM. You will need to modify your logic to throw out these cases.

---

Level 1: Slide 11 Resources: Slide 31

# Intro Monte Carlo Lab Resources

## Monte Carlo Simulation of DDM Resources

1. Slides - Monte Carlo Simulation

---

Level 1: Slide 11 Level 2: Slide 30

# Monte Carlo Python Lab, Level 2

## Monte Carlo Simulation of Python Models, Level 2

1. Continue from the first lab exercise. Now you are also concerned that your assembly line will not be as efficient and so the number of phones per machine will be lower. So draw that from a normal distribution with mean 100,000 and standard deviation of 20,000.
2. As you run the model, also store what were the interest and number of phones corresponding to the NPV. You want to see which has a greater impact on the NPV: interest or number of phones. Visualize the relationship between interest and NPV, and the relationship between number of phones and NPV. Also run a regression to quantitatively determine which has a greater effect.

---

Level 1: Slide 23 Resources: Slide 33

# Monte Carlo Python Lab Resources

## Monte Carlo Simulation of Python Models Resources

1. Slides - Monte Carlo Simulation

---

Level 1: Slide 23 Level 2: Slide 32

# Monte Carlo Excel Lab, Level 2

## Monte Carlo Simulation of Excel Models, Level 2

1. Continue from the first lab exercise. Now you are also concerned that there is varying quality in the machines, so they may have a different lifespan. Draw that from a normal distribution with mean 10 years and standard deviation of 2 years.
2. As you run the model, also store what were the number of phones and machine life corresponding to the NPV, all in Excel. You want to see which has a greater impact on the NPV: number of phones or machine life. Visualize the relationship between number of phones and NPV, and the relationship between beginning machine life and NPV. Try to determine which has a greater effect.

---

Level 1: Slide 28 Resources: Slide 35

# Monte Carlo Excel Lab Resources

## Monte Carlo Simulation of Excel Models Resources

1. Slides - Monte Carlo Simulation

---

Level 1: Slide 28 Level 2: Slide 34