Stochastic Modeling

# What is it?

Related to financial planning, stochastic modeling is a broad term that describes various methods used to simulate ranges of outcomes for systems that are to some extent random and unpredictable. More formally, the behavior of these systems is said to be non-deterministic, in that a past or present state does not fully determine its next state, as opposed to deterministic which describes straight-line,[[1]](#endnote-1) one-pass[[2]](#endnote-2) linear modeling.

There are a variety of stochastic modeling methods. The astute planner will learn the various methods and the usefulness of each depending on the characteristics of the inputs and the exact question to be answered. This critical thinking will result in the most useful results and prevent possible errors with the serious consequence of misleading clients.

Independent versus Dependent

Stochastic modeling can be broadly classified by the relational (in)dependence of the input variables in the system. First appearance is usually not enough to accurately determine the relationship of various factors; they must be closely examined.

A Markov chain is a system where future outcomes are only dependent on the current state and not on any past events. An example of this would be a game using dice like Chutes and Ladders. No matter how a player arrived at his current spot, the outcome of the next event (roll) is only dependent on where the player currently resides. Nothing can happen to the player in the next turn beyond what one roll of the dice can force to happen. If two players both reside on the same spot and arrived by different routes, they have equal chance of various outcomes.

A Bernoulli process is a subset of Markov chains that exists with only two possible results per trial that are completely independent. Variables are fully independent when current states (in addition to past states of Markov) have no predictive value for future states. If a quarter is flipped and lands on heads 9 times in a row, the probability of it landing on tails next time is still only 50 percent instead of something closer to 90 percent if it were dependent on past states.

A third category has random events with variables that are conditionally dependent. How is an event random if the future is dependent on past occurrences? The best example of this is related to the namesake of one of the most popular methods of stochastic optimization, Monte Carlo. If a person is playing blackjack and the dealer has

Stochastic Modeling

a shoe of six decks of cards randomly shuffled, the first card drawn is the only completely unknown card of the 312 possible cards. If the first card drawn is a Jack of Clubs, there is no absolute certainty of what the second card will be, but there are only five Jack of Clubs left and six of every other card. As the game progresses, the data set of inputs grows and the ability to estimate what card will be drawn increases until the very last card is fully known before being drawn.

Monte Carlo Simulation

Monte Carlo simulation is by far the most widely used stochastic process in the financial planning and investment industries, likely because:

• The simplicity of input data compared to most other methods.

• Monte Carlo may or may not consider future states as conditionally dependent to past events and current state.

• Even though events are still considered random (as not being predictable), Monte Carlo does not consider them complete chance.

• The output is generally viewed as easily understandable.

When using Monte Carlo for portfolio optimization, most people use standard deviation as the mechanism to generate random future states (usually investment returns). Because of using standard deviation, Monte Carlo assumes returns are normally distributed around the mean (average) return. This hypothesis is frequently debated in research. A normal distribution is characterized by the following criteria:

1. The majority of returns are close to the mean with a few outliers.

2. The occurrence of returns above and below the mean return would be equal.

3. Each side is a mirror image of the other divided in half by the mean (average return).

Stochastic Modeling



Figure 1 is an over-simplified histogram of how the U.S. stock market might look if it were normally distributed. A histogram visually depicts standard deviation by charting the frequency of past events and their distance from the mean (assumed to be 10% here). In the 66 theoretical events shown, returns were assumed to be between 8 and 12 percent eight times (or 12% of the time). In a normal distribution, 68 percent of the returns are within 1 standard deviation of the mean, which hints that standard deviation in Figure 7.1 is about 12 percent (between -2% and 22%). This simplified chart allows a manual illustration of how Monte Carlo works with standard deviation for portfolio optimization.

*Example:*Assume the histogram in Figure 1 is used to simulate possible returns for stocks over a period of 10 years.

With 66 events from Figure 1, assume there are 66 blank lottery balls. On 4 balls there are numbers between negative eight and negative four and on 6 balls there are numbers between negative four and zero, until all 66 possible returns are accounted for matching Figure 1 on a ball.[[3]](#endnote-3) Then, the balls are placed into the machine and one ball is drawn and the return number written down. That ball is then placed back in the machine. Then, another ball is drawn and that return number is

Stochastic Modeling

written down. This is continued as long as desired. In theory, the longer this is done, the more accurate the results will be. Obviously, at some point the incremental gain in accuracy is no longer worth the effort.[[4]](#endnote-4)

If 10 year returns are being modeled, then a reasonable strategy in this manual illustration might be to pull 10 sets of 10 years of returns. Then show the ranges on the high side and the low side and calculate the mean of the 10 return sequences. In this simplifiedcase then, 100 returns might need to be drawn to get a meaningful data set (10 sets that each have 10 years of returns).

For ease of illustration, assume two return sequences as in Figure 2.

Starting value of $100,000 with no contributions or withdrawals.

Scenario 1

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Year | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| ROR | 14% | 18% | -6% | 16% | 13% | -11% | 28% | 0% | 18% | 10% |
| Beginning | 100,000 | 114,000 | 134,520 | 126,449 | 146,681 | 165,749 | 147,517 | 188,821 | 188,821 | 222,809 |
| After Return | 114,000 | 134,520 | 126,449 | 146,681 | 165,749 | 147,517 | 188,821 | 188,821 | 222,809 | 245,090 |
| Average Return | 10.00% |  |  |  |  |  |  |  |  |
| Standard Deviation | 12.06% |  |  |  |  |  |  |  |  |

Scenario 2

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Year | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| ROR | 10% | 18% | 0% | 28% | -11% | 13% | 16% | -6% | 18% | 14% |
| Beginning | 100,000 | 110,000 | 129,800 | 129,800 | 166,144 | 147,868 | 167,091 | 193,826 | 182,196 | 214,991 |
| After Return | 110,000 | 129,800 | 129,800 | 166,144 | 147,868 | 167,091 | 193,826 | 182,196 | 214,991 | 245,090 |
| Average Return | 10.00% |  |  |  |  |  |  |  |  |
| Standard Deviation | 12.06% |  |  |  |  |  |  |  |  |

Theoretically, the numbers in Figure 2 were created through trial and error to match the return and estimated standard deviation from Figure 1. Nothing really ground breaking can be taken away from these two examples. [For later illustrative purposes, Scenario 2 in Figure 2 has been intentionally made to be a reverse sequence of the series of returns in

Stochastic Modeling

Scenario 1 in Figure 2. As a result, the average return, the standard deviation, and the final value are the same for both scenarios where there are no contributions or withdrawals.]

However, some things change when cash flows are added to the scenarios, see Figure 3.

Starting value of $100,000 with $10,000 withdrawn at the end of each year.

Scenario 1

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Year | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| ROR | 14% | 18% | -6% | 16% | 13% | -11% | 28% | 0% | 18% | 10% |
| Beginning | 100,000 | 104,000 | 112,720 | 95,957 | 101,310 | 104,480 | 82,987 | 96,224 | 86,224 | 91,744 |
| After Return | 114,000 | 122,720 | 105,957 | 111,310 | 114,480 | 92,987 | 106,224 | 96,224 | 101,744 | 100,919 |
| After W/D | 104,000 | 112,720 | 95,957 | 101,310 | 104,480 | 82,987 | 96,224 | 86,224 | 91,744 | 90,919 |
| Average Return | 10.00% |  |  |  |  |  |  |  |  |
| Standard Deviation | 12.06% |  |  |  |  |  |  |  |  |

Scenario 2

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Year | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| ROR | 10% | 18% | 0% | 28% | -11% | 13% | 16% | -6% | 18% | 14% |
| Beginning | 100,000 | 100,000 | 108,000 | 98,000 | 115,440 | 92,742 | 94,798 | 99,966 | 83,968 | 89,082 |
| After Return | 110,000 | 118,000 | 108,000 | 125,440 | 102,742 | 104,798 | 109,966 | 93,968 | 99,082 | 101,553 |
| After W/D | 100,000 | 108,000 | 98,000 | 115,440 | 92,742 | 94,798 | 99,966 | 83,968 | 89,082 | 91,553 |
| Average Return | 10.00% |  |  |  |  |  |  |  |  |
| Standard Deviation | 12.06% |  |  |  |  |  |  |  |  |

Notice, the reverse sequencing of returns does not affect the average return and standard deviation (the conventional measure of investment risk) between the scenarios in Figure 1 and Figure 2. But when those returns are applied to cash flows, the order of returns may be just as influential as the mean return over the 10 year period. If the terminal result was not impacted by return sequences, then the mean return over the horizon would be the only factor influencing the success of an investment portfolio with distributions. Yet, by using deterministic (traditional, one-pass, mean return, linear projection) financial planning, the advisor may be making the assumption (possibly unconsciously) that the sequence of returns is immaterial.

Stochastic Modeling

Multiple Assets

Monte Carlo simulation is generally much more complicated than the simple illustration shown here. Monte Carlo simulation can become so complicated so as to become impractical to illustrate manually.

One complexity that can be added to Monte Carlo simulation is the correlation of each asset with every other asset. With just the addition of one asset, the complexity of how a program generates random returns may be exponentially increased by adding the correlation interrelationships of each of the assets to each other.

A major complexity may come from the method used by a program to calculate returns. Returns may be much more complex than those shown for illustrative purposes above; monthly returns, standard deviations, and correlations may be used. Also, a program may presume a simulation of inflation (with the option to change it to a fixed amount). By simulating inflation, each monthly cash flow of every trial is randomly generated for multiple outcomes.

Roughly how many calculations might a standard simulation perform?

*Example:* Assume a 10 year (120 months) simulation, using monthly returns, invested in 5 asset classes, and adjusting monthly cash flows for inflation, and 500 trials.

Inflation simulation of cash flows:

120 months x 500 trials x correlation with 5 assets = 300,000

Calculation of investment returns:

120 months x 500 trials x correlation with 4 factors x 5 asset classes = 1,200,000

Total (not including simulation compounding return calculations) 1,500,000

In reality the calculations for generating random returns for each asset class and inflation were left out, so the 1.5 million may be dramatically underestimated.

Limitations

Correlation

The most often cited shortcoming of Monte Carlo when applied to portfolio optimization is the inability to properly reflect issues of correlation. In general, correlations are

Stochastic Modeling

assumed to be fixed (once determined), rather than being stochastic processes themselves.

Normal Distribution

As stated earlier, Monte Carlo generally assumes that returns are normally distributed and using standard deviation as an input generates random returns consistent with the probabilities of that normal distribution. If this assumption is not accurate, then the random returns based on a normal distribution could produce misleading results, by inappropriately understating the likelihood of some returns, and overstating the likelihood of others.

Historical Simulations

Historical simulations use actual historical return data to generate iterations of returns to model potential retirement scenarios.

Historical simulation methods may compensate for the serial correlation issues of Monte Carlo by linking historic periods of returns together to reflect how an asset’s returns over time do actually affect the likelihood of subsequent returns. Unfortunately, historical simulations can only reflect the way serial correlations have actually behaved in the past (which may or may not be representative of how they behave in the future), yet this may still be an advantage. It is possible serial correlations will change differently in the future, but at least some consideration is given to change.

Similarly, historical simulations may overcome Monte Carlo’s limitation concerning cross-correlations. The historic slices of data are taken across assets for the same time period, so a very true reflection of how cross-correlations have changed and will continue to change over time is factored in. Again, this may not be how cross-correlations occur in the future, but the fact that historical simulations do account for them changing over time may be useful.

Independent Discrete Historical

Independent historical simulations are the most restrictive of the three options for historical simulations. Each set of data is only able to be used one time, and the years are not overlapped or used again. Thus, if 100 years of data is available, and the

Stochastic Modeling

advisor desires to model a 20 year time frame, only 5 distinct trials are calculated. This method, with the limited data set, negates most of the benefit available from historical modeling.

120 months x 500 trials x correlation with 5 assets = 300,000

Rolling Historical

To generate a meaningful data set, rolling historical returns use a return multiple times in overlapping samples. In the prior illustration, assuming 100 yearly returns, observations would be as follows:

120 months x 500 trials x correlation with 4 factors x 5 asset classes = 1,200,000

Total (not including simulation compounding return calculations) 1,500,000

The obvious benefit is that instead of having only 5 trials to the simulation, now 81 observations are available for a 100 year time period.

Linked Historical

Linked historical simulations are a further development of rolling historical returns for the purpose of generating even more return observations. Notice in the rolling historical example the return data for year 1 is only used in the very first observation. Similarly, the return from year 100 will only be used one time in the very last observation. Other year’s returns throughout the middle of the time period will be used 20 times. Linked historical simulations generate 100 observations from 100 years of data. Take the returns from year 82 through 100 (19 years) and add year 1 to the end to make a complete 20 year return set. This will continue on until all years have been included in 20 different observations (Years 83 to 100 linked with years 1 and 2). A drawback to this method may be that it potentially creates artificial serial correlations because it connects data consecutively that wasn’t actually consecutive (e.g., years 100 and 1 will be repeatedly paired).

Limitations

The obvious concern of historical simulations is noted by the enhancements that have been developed. The number of runs available when using historical data may be very small. Even when using linked historical simulations, the observations do not even begin to match the thousands of trials most experts recommend using for Monte Carlo simulation to minimize statistical errors.

Stochastic Modeling

Conclusion

To preserve serial and cross-correlations along with non-normal distributions, modeling based on historical returns may be useful.

Bootstrap Simulations

Bootstrap simulations attempt to meet Monte Carlo and historical simulations in the middle. Historical return data is used for the return inputs. But instead of series of returns taken sequentially in the order actually observed, bootstrapping takes the returns for all assets in the same period and randomly blends those with returns from other periods.

So one iteration is the full set of returns for each asset from one year blended with other year’s returns randomly selected:

year 82, year 41, year 15, year 2, year 64, year 55, year 49, year 32, year 77, year 26…

…for 20 years then used to run one trial.

Similar to millions of possible lottery combinations from a much smaller number sequence, bootstrap simulation takes advantage of the existing data to generate returns, while allowing for limitless combinations similar to Monte Carlo.

Limitations

Bootstrapping’s compromise between Monte Carlo and historical simulations may come at the expense of the best benefits from each method.

Compared to historical simulations, the benefit of accounting for cross-correlation is retained (because returns across assets are taken from the same time period observation), but asset specific serial correlation is lost, as the time sequence of returns is randomized.

When compared with Monte Carlo, bootstrapping maintains the ability to generate practically limitless data sets, reducing the chance of sample errors due to a small sample size, but at a tremendous cost. First, the ability to make adjustments to the inputs may be lost. Second, when adding an asset with limited historical return data, the returns of all other assets will be cut down to the shortest common denominator. With Monte Carlo, a new asset can be added as long as some estimate can be made for the data points for mean, standard deviation, and correlation with other assets.

Stochastic Modeling

Conclusion

As bootstrapping is a blend of both sides of Monte Carlo and historical simulation, does a situation exist where the procedure is very useful? Not likely. Monte Carlo is the most robust and yet flexible of the three options noted. To consider the impact of serial and cross correlations after Monte Carlo, some form of historical simulation should be performed (likely rolling or linked), remembering the possibility of statistical errors from the smaller input sample size. Between these two stochastic processes, the informed advisor should be able to reasonably test almost any desired course of action. Obviously with the computing capacity available today, the few extra minutes required to use each option is worth the insight that might be gained, in the right situation.

# When is the use of Stochastic Modeling Indicated?

Stochastic modeling benefits retirees by identifying simulated outcomes divided into frequency of occurrence in relation to other outcomes.

When working with a client, many planners have surprising success when using stochastic modeling interactively in meetings to assess opportunity costs and maximize utility, even if the client is unable to easily define or quantify utility. Frequently, a new financial planning client is looking for the one answer to the question, “When can I retire?” Even if the retiree fully understands what they want in retirement in relation to what they are willing to give up, they typically will not volunteer that information; however, modeling helps draw it out by giving them control over flexible inputs and “What if” scenarios.

Engaging these two factors (mutual discovery and multiple scenario output) with stochastic modeling may increase client allegiance to the resulting plan.

# What are the Requirements?

Working with a client using stochastic modeling, requires a great deal of skill and practice. One of the most important considerations is the output. Generally, output is best represented graphically with charts and summary tables, with other tables and detail available behind it. When using stochastic modeling interactively with a client the output must be easily understandable; it may help to have an available explanatory page in a print out that can be given to the client. The explanatory page must have appropriate disclosures, input assumptions for the given output, and absolutely must state something about how stochastic modeling is not a predictor of the future. If a client

Stochastic Modeling

asks the advisor’s opinion on the likely accuracy of the output, a quick walk through the background data generated may help the client accept the validity of the illustration.

The advisor must be able to quickly show the detailed information behind each observation (iteration). Reviewing detailed output of survival and failure of various iterations (with good explanation from the advisor) that succeed and fail differently than average return would predict may validate the concern some clients may have with straight-line one-pass linear projections.

From the planner’s point of view, if working interactively with a client, the computer program must have easy and quick navigation for on-the-fly changes and be able to run trials and return results in a timely manner. Ideally, possible input errors would be easily discovered and allow prompt modification. Any internet-based program should be tested for ease and speed in a trial run before attempting to interact during a client meeting. Too much waiting or software hang-ups and the client might lose confidence in the output as the perceived comfort the advisor has with the program is transposed onto the genuine validity of the program.

To help clients feel at ease with a very technical concept, developing a standard preamble to the actual exercise is very useful. If the advisor truly sees value in the process, this should resonate with clients enough to keep even completely uninterested clients mildly engaged.

Finally, having a few article reprints or information on further references (web sites, history of development) is very helpful to lend credibility to the process. Even if the client does not choose to read them, the chance of increasing receptivity is worth the effort.

# Where and how do I get it?

***Decisioneering*** www.crystalball.com

Crystal Ball is an Excel based program that is very customizable and used across broad applications by engineers, scientists, and mathematicians. A financial planning version is available. A trial version can be downloaded.

***Financeware*** www.financeware.com

Online application with Monte Carlo, historical, and bootstrap available. AASim is available as a desktop application.

Stochastic Modeling

***Financial Engines*** www.financialengines.com

Founded by William F. Sharpe, Financial Engines (online) has become widely used for retirement plan participants with $7 billion that it advises to. A financial institutions version is available.

***MoneyGuidePro*** www.moneyguidepro.com

The Stress Testing component of MoneyGuidePro allows Monte Carlo and historical rolling returns in numerous capacities. A detailed explanation of the calculations is available at www.moneyguidepro.com/Pdf/Help/StressTestingCalculationDetails.pdf. In some cases, the average portfolio rate of return is used for calculations instead of individual asset classes.

***Money Tree*** www.moneytree.com

Comprehensive financial planning program. Only Monte Carlo simulation is available, but modifying return assumptions is allowed. Money Tree seems to only calculate Monte Carlo based on one weighted rate of return.

***NaviPlan*** www.naviplan.com

With EISI’s (parent company) acquisition of Financial Profiles, Inc. the combined firm will have 12,000 users and will be the dominate provider of third-party planning software in the industry. The comprehensive NaviPlan software has a web based application and two desktop versions (standard and extended). Both only offer Monte Carlo simulation among other features.

***Profiles*** www.profiles.com

Profiles + Professional has Monte Carlo as an optional feature. Profiles + has its origin in the insurance arena and was recently purchased by EISI, the owners of NaviPlan.

As each is quickly changing their programs, the most instructive method of review is for the advisor to experiment with a trial version. Then take a client situation and enter it into multiple versions.

1. Run a mock client scenario.

2. Run the same scenario a second time. Observe how results change.

3. Find out what is being randomized.

Stochastic Modeling

4. Change each input and see how output is affected.

a. If inflation is a static number, what does a 1 percent increase in assumed inflation cause in the output?

b. What affect does excess longevity have on the output?

# Where can I find out more about it?

1. Americks, Veres, and Warshawsky, “Making Retirement Income Last a Lifetime,” *Journal of Financial Planning*, Dec. 2001.

2. Kautt and Hopewell, “Modeling the Future,” *Journal of Financial Planning*, Oct. 2000.

3. Hopewell, “Decision Making Under Conditions of Uncertainty: A Wakeup Call for the Financial Planning Profession,” *Journal of Financial Planning*, Oct. 1997.

4. Kautt and Wieland, “Modeling the Future: The Full Monte, the Latin Hypercube and Other Curiosities,” *Journal of Financial Planning*, Dec. 2001.

5. Loeper, “Are You Modeling What You Intended?” *Wealthcare Capital Management White Paper,* Aug. 15, 2003.

6. Metropolis, “The Beginning of the Monte Carlo Method,” *Los Alamos Science,* Special Issue 1987.

7. Merton and Samuelson, “Fallacy of the Log-Normal Approximation to Optimal Portfolio Decision Making over Many Periods,” *Journal of Financial Economics,* May 1974.

8. Bengen, “Conserving Client Portfolios During Retirement, Part IV,” *Journal of Financial Planning*, May 2001.

9. Guyton and Klinger, “Decision Rules and Maximum Initial Withdrawal Rates,” *Journal of Financial Planning,* May 2006.

1. Typically, a set rate of return with possible adjustments when investment allocations change. [↑](#endnote-ref-1)
2. The financial plan has only one shown outcome. [↑](#endnote-ref-2)
3. In actual practice, a computer randomly generates specific numbers to match a normal distribution, not ranges. Randomly generating numbers based on average return and standard deviation (and possibly incorporating correlations) does not use a fixed sample, such as a set number of balls. [↑](#endnote-ref-3)
4. Unfortunately, there is no yardstick for knowing when there is enough. Each simulation is situation dependent. [↑](#endnote-ref-4)