# Poppertech Calculator Forecasting Methods

**Introduction:** The purpose of forecasting is to anticipate future events such that reasonable decisions can be made in the present to plan for their possible consequences. The Poppertech Calculator shows how investors can anticipate their ability to reach their individual financial goals given their present decisions regarding investment allocation. Two main types of forecasts exist: deterministic and probabilistic. The Calculator employs both.

Deterministic forecasts consist of certain statements regarding future events:

* “Exxon’s stock will trade between $90 and $100 per share at the end of this year.”
* “Mabel requires $65,000 over the next 10 years.”

In finance, deterministic forecasts are used to both plan and assess the viability of complicated transactions, such as buying a new piece of equipment for a factory.

In contrast, probabilistic forecasts consist of uncertain statements regarding future events:

* “There is a 60% probability of Exxon’s stock trading between $90 and $100.”

Typically, when there is some level of control over an outcome, deterministic forecasts predominate. In the examples above, Mabel has some degree of control over her expenses. Therefore, a deterministic forecast seems appropriate. Correspondingly, the Calculator uses deterministic forecasts for an investor’s required cash flows.

In contrast, probabilistic forecasts predominate when no person or organization controls an outcome. For example, no person or organization controls Exxon’s share price. It is determined by a confluence of factors. Consequently, a probabilistic prediction is more suitable, and the Calculator implements probabilistic investment forecasts.

**Deterministic Versus Probabilistic Forecasts:** Deterministic forecasts are both easier to understand and more common than probabilistic forecasts. Probabilistic forecasts are difficult to understand in part because the meaning of probability itself is the subject of a contentious debate. When an event has a 50% probability of occurring, does it mean: 1) if the event were replayed a large number of times, we would see that the event occurred half of the time; or 2) the person who is making the forecast is uncertain about whether the event will occur or not? Professional statisticians have debated this question for decades, so it is not surprising that many people find probability to be confusing. In public discourse, weather forecasts are among the few examples probabilistic predictions.

If probabilistic forecasts are both uncommon and complex, then why are they useful? Why do weather forecasters predict a 60% chance of rain? Probability conveys the uncertainty of a forecast. Alternatively, It expresses the confidence of the forecaster. When it is impossible to control an event and knowledge regarding its outcome is necessarily limited, probabilistic forecasts capture these limitations mathematically. Since human knowledge about the future is inherently limited, probabilistic forecasts are much more realistic than deterministic ones.

Therefore, a tradeoff exists: deterministic forecasts are easier to communicate but less realistic than probabilistic forecasts. The Calculator implements deterministic forecasts of the investor’s required cash flows because this facilitates communication between the financial advisor and investor. However, it employs probabilistic forecasts for the investments because the realism of these forecasts is best aligned with the investor’s interests.

**Understanding the Calculator’s Forecasts for a Single Investment**

As discussed above, the required cash flow forecasts are deterministic and standard. They mimic the spreadsheet models that financial advisors currently use to determine cash needs. However, the investment forecasts are probabilistic and are much less common. Therefore, they require further explanation.

An analogy with weather forecasts may aid understanding why the investment forecasts need to be complex and realistic:

*Scenario 1: Suppose you want to know whether to bring an umbrella to work. You check the weather forecast, which says there is a 60% chance of rain. Since rain is likely, you decide to carry an umbrella.*

In Scenario 1, the decision is binary and simple: either bring or do not bring an umbrella. In parallel, the forecast mirrors the decision: a 0% probability of rain indicates it will certainly not rain and no umbrella is required; whereas, a 100% probability of rain indicates it will certainly rain and an umbrella is required. However, a more complex decision would more complex forecasts:

*Scenario 2: A school administrator is deciding whether to cancel school. If there is less than one inch, school should stay open. If there is more three inches, school should be cancelled. If there is one to three inches, the administrator must make a judgment call based on the timing and likelihood of the snow. Since the weatherman forecasts 80% chance of 1-3 inches of snow starting in the early morning, she decides to close school.*

The outcomes in Scenario 2 are complex, and the forecast must match this complexity to be useful. Rather than a simple binary outcome, multiple scenarios are possible. The forecaster must express his confidence regarding intervals of outcomes to aid the administrator’s decision.

Analogous to Scenario 2, investment decisions are both uncertain and complex. Consequently, they need forecasts that adequately match them. The Calculator requires forecasts that area specified similarly to Scenario 2 for the value of a $100 investment after one year. Each forecast consists of four intervals and these are named: “Left Tail,” “Left Normal,” “Right Normal,” and “Right Tail.” These terms have the following definitions:

* “Left Tail”: the bottom outcomes whose combined probability totals 10%
* “Right Tail”: the top outcomes whose combined probability totals 10%
* “Normal”: the middle outcomes whose combined probability totals 80%
  + “Left Normal”: the interval from the “Left Tail” to the most likely outcome
  + “Right Normal”: the interval from the most likely outcome to the “Right Tail”

The Calculator requires inputs from a forecaster to determine each of these intervals. To elicit these inputs, the forecaster should ask the following questions:

* What interval can I say with 80% confidence will contain the value of a $100 investment over one year?
* What do I think the most likely result is?
* What is the worst possible outcome?
* What is the best outcome?

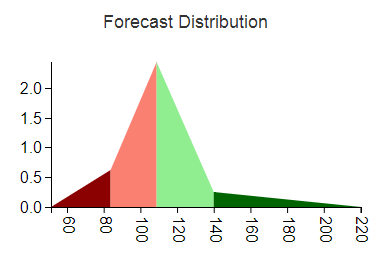
Based on the answers to these questions, there are five inputs to each forecast:

1. The minimum possible outcome (“Minimum”)
   1. Lower bound of the Left Tail
2. The point at the bottom of the 80% confidence interval (“Worst Case”)
   1. The probabilities of the outcomes below this point sum to 10%
   2. Upper bound of the Left Tail
3. The most likely outcome (“Likely”)
   1. The highest point on the Forecast Distribution Graph
   2. Divides the “Normal” region into “Left Normal” and “Right Normal”
4. The point at the top of the top of the 80% confidence interval (“Best Case”)
   1. The probabilities of the outcomes above this point sum to 10%
   2. Lower bond of the Right Tail
5. The maximum possible outcome (“Maximum”)
   1. Upper bound of the Right Tail

The Calculator determines the Forecast Distribution and investment simulations based on these inputs. To derive the Forecast Distribution from the inputs, a method is required for estimating probabilities for points between inputs. Therefore, the Calculator incorporates the following assumptions:

* The Left Tail consists of a triangle whose area equals 10% probability
  + One side spans the interval from the Minimum to the Worst Case inputs along the x-axis
  + One side is vertical at the Worst Case
  + One side connects the previous two segments
* The Right Tail consists of a triangle whose area equals 10% probability
  + One side spans the interval from the Maximum to the Best Case inputs along the x-axis
  + One side is vertical at the Best Case
  + One side connects the previous two segments
* The Left Normal region consists of a trapezoid
  + One side spans the interval from the Worst Case to Likely inputs along the x-axis
  + One side is vertical at the Worst Case and is shared with the Left Tail triangle
  + One side is vertical at the Likely input and is shared with the Right Normal region
  + One side connects the Worst Case and Likely inputs at their maximum height
* The Right Normal region consists of a trapezoid
  + One side spans the interval from the Likely to Best Case inputs along the x-axis
  + One side is vertical at the Best Case and is shared with the Right Tail triangle
  + One side is vertical at the Likely input and is shared with the Left Normal trapezoid
  + One side connects the Likely and Best Case inputs at their maximum height

The Calculator determines the heights in the Forecast Distribution based on these assumptions and the probabilities associated with each region. Please see Appendix 1 for the calculation formulas and mathematical formulation of the Forecast Distribution.



**Determining Reasonable Values for Forecast Inputs**

As described in “Setting Calculator Defaults,” the Calculator sets the default inputs based on historical data. The forecaster then uses these defaults as a baseline when changing the inputs. To help guide these changes, the Calculator displays the Mean, Standard Deviation, Skewness, and Kurtosis of the Forecast Distribution. The Mean measures what the forecaster expects to occur on average. Setting the Mean to be less than the defaults signifies the forecaster expects the investment to perform poorer on average than the historical period. Standard Deviation measures dispersion in future performance. All else equal higher dispersion, or volatility, results in a lower chance the investment will enhance an investor’s ability to reach a goal. Skewness measures asymmetry of future performance. An investment that performs well 99% of the time may still be undesirable if 1% of the time it results in a catastrophic loss, and negative skewness captures such asymmetrical payouts. Finally, Kurtosis measures the weight of the tails, or the possibility of large magnitude payouts in either direction. All else equal, Kurtosis increases the risk of the investor missing the goal. Therefore, Standard Deviation, Skewness, and Kurtosis all relate to risk.

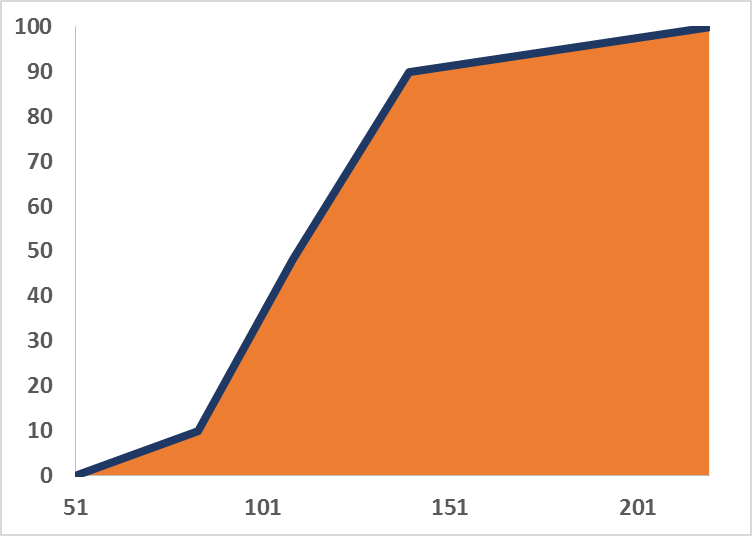
**Mapping Uniform Random Numbers to Simulations Using the Forecast Distribution**

Thus far, the discussion has focused on forecasting investment performance. However, the underlying goal is to simulate investment performance, and subsequently, determine the probability of an investor reaching a cash flow goal based on these simulations. Therefore, a method of converting the forecasts to simulations is required.

Before it is possible to map forecasts to simulations, an intermediate step is required. It is necessary to determine the probability associated with each point on the Forecast Distribution. As shown in the above sections, the Forecast Distribution represents probabilities as areas underneath. Therefore, the Calculator determines the areas underneath and to the left of each point of the Forecast Distribution. This is called the Cumulative Forecast Distribution. Figure XX shows the results graphically. Appendix 2 derives the resulting mathematical formulas.

Next, the Calculator needs random numbers that it will transform to investment simulations using the Cumulative Forecast Distribution. Many computing languages supply pseudo-random numbers on the interval from 0 to 1 such that the probability of drawing any given number is equal. The underlying distribution for such a random number generator is known as the Uniform Distribution.

Since each random number drawn from the Uniform Distribution is on the interval from 0 to 1, each draw shares similar characteristics with a probability. The Cumulative Distribution plots probabilities on its y-axis. Consequently, each “probability” drawn maps to a horizontal line drawn across the y-axis. The intersection of this horizontal line and the Cumulative Distribution determines the investment simulation. The Calculator repeats these draws thousands of times to obtain a representative sample from the Cumulative Distribution. Formally, the general type of mathematical modeling that use numerical approximations is called Numerical Analysis. This can be contrasted with Analytical methods that solve systems of equations deterministically. The specific method employed is called Monte Carlo analysis. Numerical techniques are useful when equations cannot be solved analytically or when the Analytical solutions are extremely complicated. As the next section discusses, the modeling of investment portfolios requires forecasting the connections between investments. This step increases both the complexity of the forecasts and the benefits from using Monte Carlo analysis.

****

**Understanding the Forecasts for Multiple Investments**

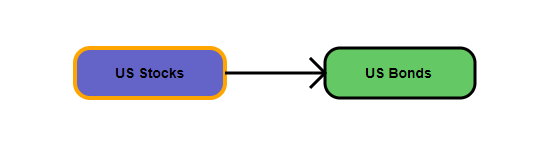
Thus far, the discussion has been focused on forecasting performance of a single investment. However, simulations of investment portfolios are required to calculate the probability of an investor reaching a goal. This adds a layer of complexity because the Calculator needs to account for the relationships between investments as well as their individual performance. These relationships combined with individual performance determine the portfolio’s risk. A contrived example may help illustrate this last point. Suppose that a $100 investment in both Exxon and General Electric and each can only gain $10 or lose $5 in any given year.

* Scenario 1: If both Exxon and GE always increase or decline simultaneously, then the investor would either make $20 or lose $10 per year.
* Scenario 2: If Exxon always increases while GE declines and vice-versa, then the investor would always make $5.

Clearly, Scenario 2 has much lower risk than Scenario 1 and shows why the connections between investments matter.

**Causal model between investments**

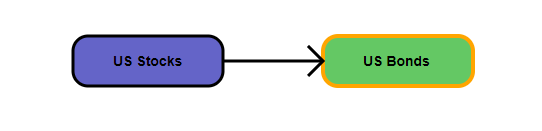
The Calculator uses a Bayesian network to relate investments to one another. Figure XX diagrams the relationship between the two potential investments used in the Calculator: stocks and bonds. Bayesian networks are a type of Directed Acyclic Graph. “Directed” means that causal relationships are identified: the performance of stocks affects bonds. Stocks are the “parent” of bonds, and bonds are the “child” of stocks. The arrow in the diagram indicates the causal direction. “Acyclic” means that the causal relationships only flow in one direction; arrows cannot have two heads, and no child of bonds could have an arrow directed towards stocks. As an aside, the Calculator could have just easily assumed the opposite causal direction: bonds affecting stock prices. In finance, US stock prices are a standard causal factor; consequently, the Calculator is follows convention.



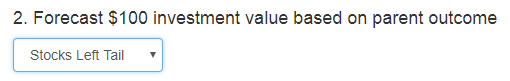
**Forecasts for Multiple Investments**

The Calculator simulates the performance of multiple investments similarly to a Single Investment with a single crucial difference: each Forecast Distribution for the child investment depends on the result of the parent. In particular, different Forecast Distributions are required for outcomes falling in the parent Left Tail, Left Normal, Right Normal, and Right Tail regions. In other words, the child Forecast Distributions are “conditional” upon the parent outcome. Therefore, the forecaster must specify the Minimum, Worst Case, Most Likely, Best Case, and Maximum outcomes for each of the four parent regions, or a total of 20 parameters per child. There is a tradeoff between forecaster time and simulation specificity, and the design of the Calculator attempts to balance this appropriately.

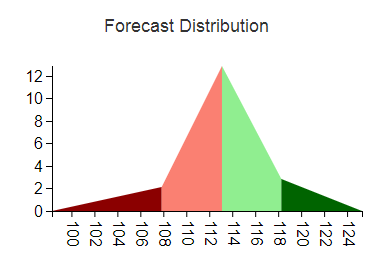
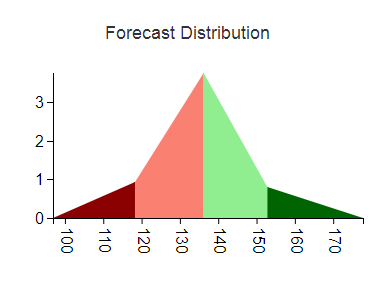
An example may help to illustrate how the Calculator simulates the performance of child variables. Figure YY depicts the network after the user selects bonds as the variable to forecast.

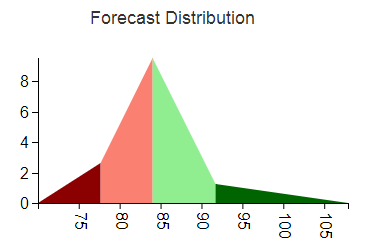
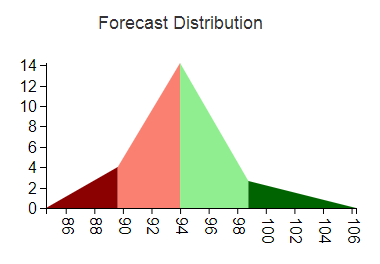


Next, the user selects the parent region. Figure ZZ shows that the user has selected “Stocks Left Tail” as the parent region. Since a different set of forecasts corresponds with each parent region, the forecast inputs and Forecast Distribution change each time the user selects a different value from this dropdown. Figure AA shows the default Forecast Distributions for each parent region.



***Forecast Distributions for US Bonds conditional on a US Stock outcome in the a) Left Tail, b) Left Normal, c) Right Normal, d) Right Tail***

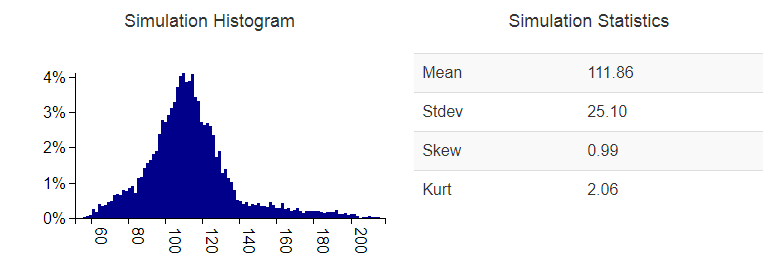




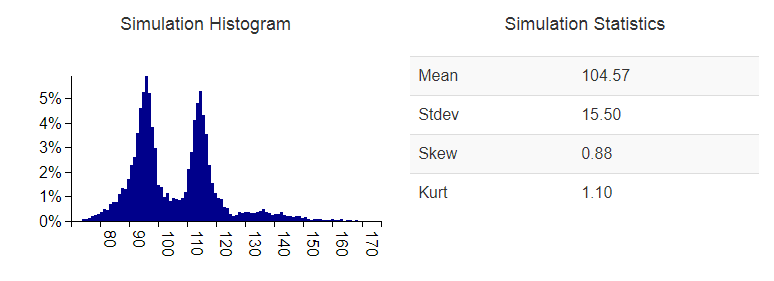
**Simulating Multiple Investment Outcomes**

Simulations for multiple investments parallel the forecasts. First, the Calculator simulates the top level variable using the same steps as described in the section “Mapping Uniform Random Numbers to Simulations Using the Forecast Distribution.” Then, it the appropriate Forecast Distribution for the child based on the outcome. For example, suppose the parent drew an outcome from the Left Tail of its Forecast Distribution. Then, the Calculator would draw the child outcome from Left Tail Forecast Distribution shown in Figure AA. The above steps would be repeated for all of the simulations.

The Calculator shows the simulation results in the “Simulate Investment Results” section. Since all of the draws for the Parent come from the same Forecast Distribution, the Simulation Histogram mirrors the Forecast Distribution and has similar statistical properties. Figure BB illustrates this result.

****

In contrast, the Calculator draws simulations for the child from four different distributions. Consequently, the child Simulation Histogram does not resemble any of the input Forecast Distributions and the statistics are also different. Figure CC demonstrates this.



**Key Benefits and Conclusion**