Frequent Once-in-a-Lifetime Crises

Randal Allen, John Volpi
Lone Star Analysis
Addison, TX
rallen@lone-star.com, jvolpi@lone-star.com

ABSTRACT

Nassim Taleb (Taleb 2010) labels black swans with the following attributes: rarity, impact, and retrospective apparent predictability. However, Benoit Mandlebrot (Wright 2007) claims gray swans are events of considerable nature, which are predictable and for which one can take precaution. Admiral Nimitz is quoted as stating, “The war with Japan had been enacted in the game rooms at the War College by so many people and in so many different ways that nothing that happened during the war was a surprise—absolutely nothing except the kamikaze tactics toward the end of the war.” Surely, this may have been a black swan from his perspective.

This paper addresses three questions posed at the “2015 I/ITSEC Black Swan Kickoff.”

1. How do we prepare, organize, train and equip for Black Swan resiliency?
2. How can Modeling and Simulation be used to analyze and prepare or create a Black Swan?
3. Can we develop complex adaptive models and simulation tools that will enable the analysis?

The authors will follow Dr. Mandelbrot’s assertion to answer the first question. For the second and third questions, we will outline a procedure to use modeling and simulation with prescriptive analytics to reduce the potential intractability of black swans, thus demoting their status to gray. With frequency histograms and curve-fitting, we first show how distributions with thin-tails don’t fully account for risk, while fat-tail distributions better fit extremely rare event data. Then, by applying Percent Point Functions and stochastic optimization techniques to a Monte Carlo simulation of fat-tailed distributions, we show which configuration of input parameters creates a black swan. Given our approach, we offer an analytics method to evaluate black swan events and downgrade them to gray swan events.

ABOUT THE AUTHORS

Randal Allen has over 20 years of industry experience and has been with Lone Star Analysis since 2006. As Chief Scientist, he is responsible for applied research and technology development. He is a certified modeling and simulation professional (NTSA). He is the co-author of “Simulation of Dynamic Systems with MATLAB and Simulink,” 2nd ed. He is also an Associate Fellow of the American Institute of Aeronautics and Astronautics (AIAA). His holds a Ph.D. in Mechanical Engineering from the University of Central Florida, an Engineer’s Degree in Aeronautical and Astronautical Engineering from Stanford University, an M.S. in Applied Mathematics and a B.S. in Engineering Physics, both from the University of Illinois (Urbana-Champaign). He also serves as an Adjunct Professor in the Mechanical and Aerospace Engineering (MAE) department at UCF.

John Volpi is the Chief Technology Officer (CTO) of Lone Star Analysis and has served in this role since 2004. He is responsible for all technical activities, intellectual property evaluation, and process development. He began his professional career at Texas Instruments as a theoretical systems analyst and evolved into Systems Engineering, where he was named a Senior Member of the Technical Staff. He attended the Illinois Institute of Technology (B.S. Physics) and Michigan State University (M.S. Physics). He also attended graduate classes at Southern Methodist University and the University of Southern California. He has over 30 patents awarded or pending. He has been a member of the IEEE for over 30 years and was named a Senior Member for his work in wireless technologies and his efforts in the development of Intellectual Property. In 2012, John was awarded the Tech Titans Award for Corporate CTO by the DFW Metroplex Technology Business Counsel out of a pool of 4,000 firms.

Frequent Once-in-a-Lifetime Crises

Randal Allen, John Volpi
Lone Star Analysis
Addison, TX
rallen@lone-star.com, jvolpi@lone-star.com

INTRODUCTION

Before addressing the three questions from the "2015 I/ITSEC Black Swan Kickoff," a little background on black swans and gray swans is in order.

A black swan is an improbable event with colossal consequences. It is a metaphor for believing something is impossible until the belief is disproven. For example, all swans were assumed to be white, and black swans were thought to be non-existent until discovered in Western Australia. In *The Black Swan,* Taleb (2010, p. xxii) defines three black swan attributes: (1) It's an outlier, outside the realm of expectation because nothing in the past points to its possibility; (2) It brings extreme impact; and (3) We concoct explanations for it after the fact, making it seem predictable. In short, the three attributes are: "rarity, impact, and retrospective apparent predictability" (Taleb 2010). Also, by symmetry, non-occurrence of a seemingly certain event is also a black swan. Furthermore, lack of evidence of black swans doesn't mean they do not exist.

While swans of unusual character are labeled black, Benoit Mandlebrot (Wright 2007) claims we can predict something of their behavior – and in doing so, they are no longer black, but can be thought of as gray. They only seem black if we fail to acknowledge their potential existence or we fail to look. The modeling and simulation tools applied to examine potential black swans are inherently stochastic, meaning they are based on probabilistic inputs and outputs are viewed from a statistical perspective. In this paper, we discuss how such tools can cause seemingly black swans to fade to gray – and in doing so, we address the second and third questions from the "2015 I/ITSEC Black Swan Kickoff" – "How can Modeling and Simulation be used to analyze and prepare or create a Black Swan?" and "Can we develop complex adaptive models and simulation tools that will enable the analysis?"

An appropriately visualized model architecture may identify the succession of events leading to the colossal black swan which may be confirmed by examination of Tornado charts (specialized bar charts) and the Percent Point Function. Additional help with exposing black swans is available through stochastic optimization, where random variables appear in the formulation of the optimization problem thus producing a random objective function for which random iterates are employed to solve the problem. When equipped with these tools and their insights, we can reduce the surprise of a black swan, rendering it gray, and thus prepare, organize, train and equip for black swan resiliency.

MODELING RARE EVENTS

With a general understanding of black and gray swans, a model is created such that rare events may be simulated. We need a distribution that shows non-zero probabilities for data lying far from the mean on either side. In this paper, we use the financial markets to make our points, because most of us can relate to risk versus reward from a financial perspective. We will see that if thin tail distributions are used, risk is modeled too conservatively; whereas a fat tail distribution exposes a greater degree of risk and the potential for a black swan. The analogy can be generalized to a portfolio of lines of business, where different business opportunities may be assessed for return on investment, given its corresponding risk.

A recent example of a black swan event is the financial collapse of 2008. While we're not interested in the cause of the collapse, per se, we are interested in one of the many lessons learned, i.e. observance of the fat tail distribution as a more accurate representation of the collapse, and thereby categorize it as a black swan event. While implementing a normal distribution in a Monte Carlo simulation is far superior to simply using average values of risk and return

(Savage 2009), the "thin tail" of the normal distribution assigns negligible probability to data far from the mean. Harry Markowitz (Markowitz 1952; Markowitz 1979; Markowitz 1999) consistently warned that distribution selection was tricky and urged that when moving from theory to practice, some caution was warranted. Benoit Mandlebrot (Mandlebrot, 1963) found price changes in some markets (especially cotton futures) were well described by Lévy stable distributions. Eugene Fama (Fama 1963) performed similar research to what is presented here and further demonstrated the merits of "fat tail" distributions in stocks. Paul Kaplan (Kaplan 2012) shows a log-stable distribution (see Appendix) captures the non-zero probability of occurrence for rare events far from the mean. The log-stable is a generalization of the log-normal distribution commonly used to model investment return. It assumes the logarithm of one plus the decimal form of risk and return following what Mandlebrot referred to as a stable Paretian distribution (Wright, 2007).

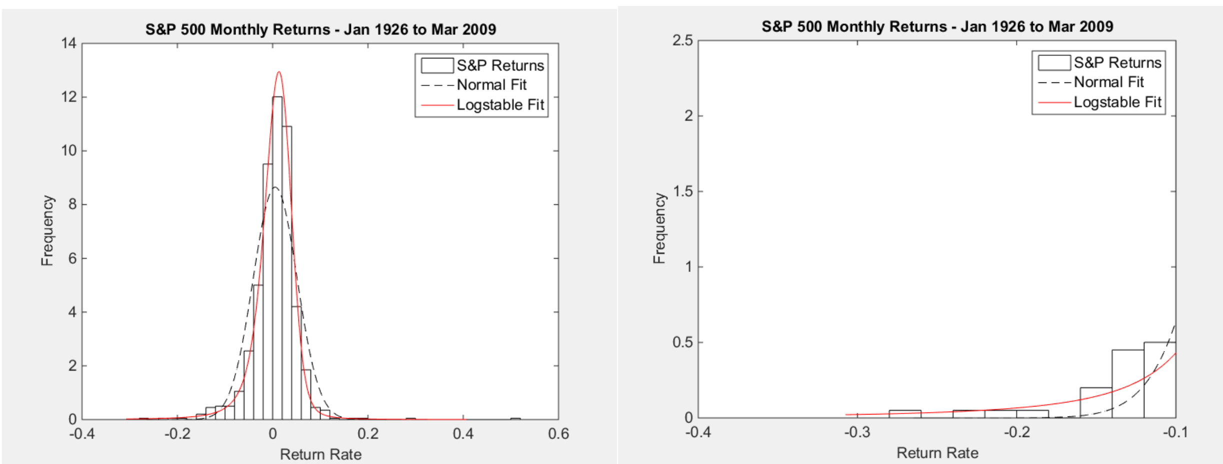

Figure 1 – S&P 500 data from January 1926 to March 2009 fitted with Normal (left) and Log-Stable (right) Distributions

The left pane of Figure 1 shows monthly returns of the S&P 500 stock index from January 1926 to March 2009 as represented by the frequency histogram. Historical returns over this time period include maximum monthly losses of 26% in November 1929, 24% in April 1932, 20% in October 2008, and 19% in December 1931; while the maximum gain was +50% in August 1932. The data is fitted with a normal distribution (dashed line) and a log-stable distribution (solid line). In the right pane, a closer examination displays the characteristics of the normal distribution's thin tail (dashed) versus the log-stable distribution's fat tail (solid). The normal distribution shows a negligible probability of losses beyond 16% (-0.16). While the theoretical tail of the normal distribution extends to infinity, it is clear from this exploded view that the probability of a 16% loss is practically zero. Using the normal distribution curve fit parameters, the actual probability is only 1.4%. Contrast this with the log-stable distribution which more accurately shows losses beyond 26% are indeed possible. In fact, using the log-stable curve fit parameters, the probability of a 16% loss is 9.6% - almost seven times more likely to occur.

From this analysis, we conclude that the log-stable distribution is superior to the normal distribution for modeling rare events of this type. While we have shown this to be true for the S&P 500, if rare events have a non-zero probability of occurrence in any practical application, the log-stable distribution should certainly be considered as the apparatus of choice.

MEASUREMENT OF PORTFOLIO RISK

In order to show the impact of portfolio risk, we model an aggressive asset allocation with the percentages shown in Table 1, where each asset class is represented by a corresponding Exchange Traded Fund (ETF) ticker symbol, e.g. small cap stocks are represented by the ETF ticker symbol IWM, etc. Historical monthly return data for each ETF was obtained from the Investools / TD Ameritrade database.

Table 1 – Portfolio Asset Allocations and ETFs

Asset Class	Allocation	ETF
Small Cap Stocks	20%	IWM
Mid Cap Stocks	15%	MDY
Large Cap Stocks	5%	SPY
Int'l Developed Stocks	5%	EFA
Int'l Emerging Stocks	10%	EEM
Corporate Bonds	15%	LQD
Government Bonds	10%	AGZ
Real Estate	10%	IYR
Commodities	10%	DBC

Furthermore, we set up two portfolios according to this asset allocation: one called the Normal Portfolio where historical monthly returns of each asset class are fitted with normal distributions and the other, called the Log-Stable Portfolio, where the historical returns are fitted with log-stable distributions. The following plots were generated from 10,000 Monte Carlo trials.

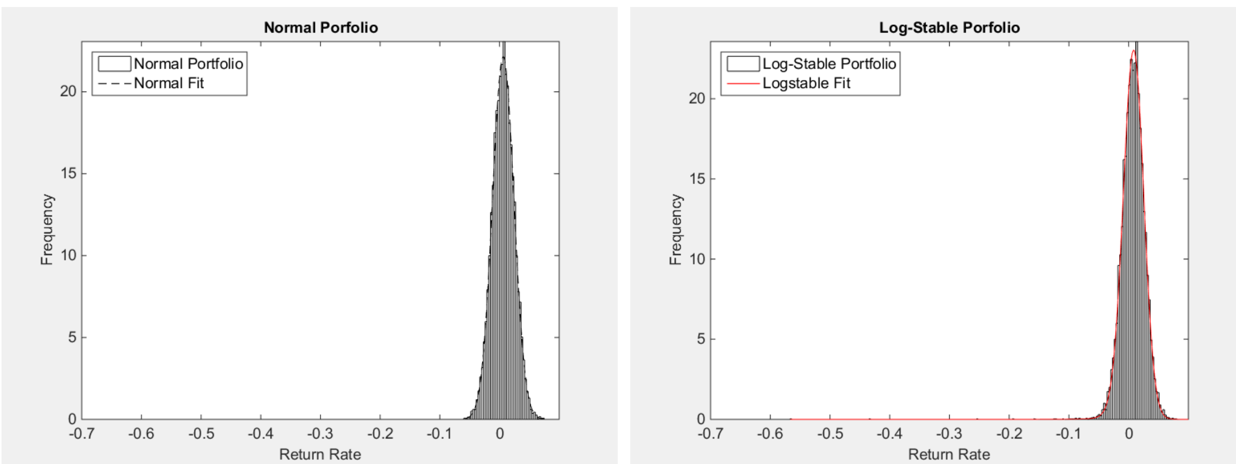

Figure 2 – Portfolios of Asset Allocations Fitted with Normal (left) and Log-Stable (right) Distributions

From Figure 2, it is difficult to tell the difference between the Normal Portfolio and the Log-Stable Portfolio.

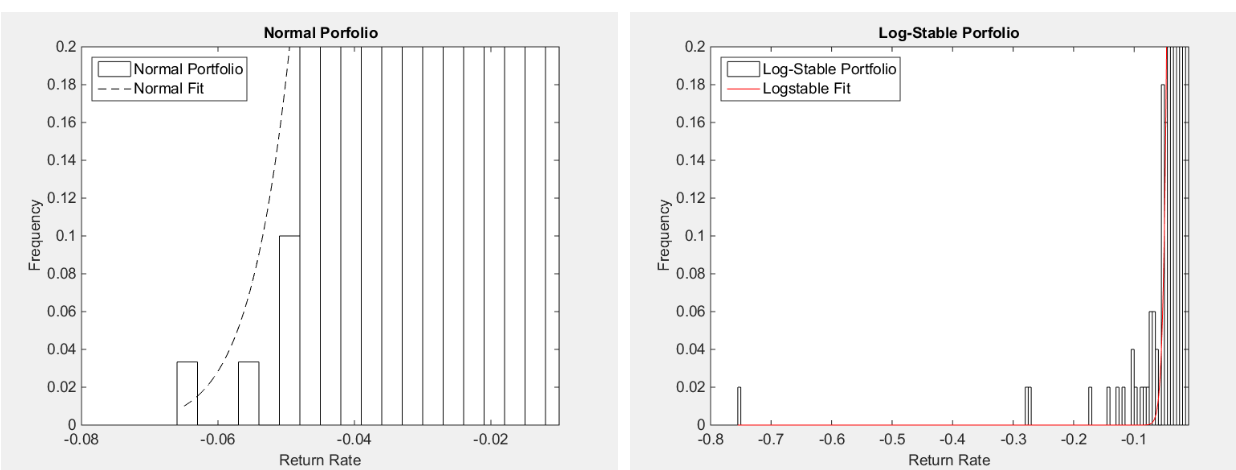

Figure 3 – Zoom of Asset Allocations Fitted with Normal (left) and Log-Stable (right) Distributions

However, upon zooming-in, we see (left pane of Figure 3) the maximum loss for the Normal Portfolio is 7% (-0.07) with a maximum gain of 7% (not shown). One may recall the normal distribution is characterized by the mean plus or minus the standard deviation and is therefore symmetric about the mean. The Log-Stable Portfolio (right pane of Figure 3) shows a maximum loss of 71% (-0.71) with a maximum gain of 16% (not shown). This left-skew (-71% versus +16%) is based on the data and the characteristic parameters of the log-stable distribution: alpha, beta, gamma, and delta (see Appendix).

For the S&P 500 data from January 1926 to March 2009, alpha = 1.5901, beta = -0.5586, gamma = 0.0219, and delta = 0.0023. Here, we see the negative value of beta as representing the left skew (see Appendix) corresponding to more risk than reward. If the data had been such that beta was positive, the distribution would have been right skewed with returns being greater than risk (an elusive investment). Incidentally, fat tails can occur on either or both sides of the distribution, depending on the data being fitted to the log-stable distribution.

Importantly, the results illustrate how the log-stable distribution more accurately predicts high levels of risk, ten-fold. Both normal and log-stable distributions are fitted to the same historical return data. Yet, the normal distribution models risk at only 7%, while the log-stable more accurately characterizes the risk at 71%, for the portfolio. This sheds a little light on the 2008 financial collapse from a risk versus reward perspective. There was actually a higher degree of risk present than otherwise indicated by normal distributions.

Again, while a financial portfolio has been used to show how the log-stable distribution is superior to the normal distribution, if rare events (e.g., earthquake magnitudes, city populations, sizes of power outages, etc.) have a non-zero probability of occurrence (either left-skewed or right-skewed), fitting the event data to a log-stable distribution will model these characteristics, raising our awareness and allowing to prepare, organize, train and equip for black swan resiliency.

EXPOSING BLACK SWANS

This is all rudimentary with simple portfolios of historical returns separated into normal and log-stable distributions. What if your model of influential architecture contains thousands of inputs, including normally distributed data as well as (rare event) log-stable distributed data?

In this section, we model a portfolio with all asset classes fitted to a normal distribution except for one asset class in order to see which tools are useful for finding which input is causing the downside risk. The tools available are Tornado charts, Percent Point Functions, and stochastic optimization methods.

Tornado Charts

Tornado charts are specialized bar charts which show how varying an input impacts the output. In our case, Tornado charts (Figures 4A and 4B below) show the impact of each asset class variation on the overall portfolios. The asset class which has the greatest impact on both portfolios is IWM or small cap stocks. Changes in the return of IWM can lower the nominal (Normal Portfolio) return by as much as 178% or raise it by 125%. Similarly, changes in the return of IWM can lower the nominal (Log-Stable Portfolio) return by as much as 164% or raise it by 127%. Small cap stocks (weighted at 20%) are more volatile. The asset class with the least impact is AGZ or government bonds. Changes in the return of AGZ have minimal impact on the nominal (Normal Portfolio) return, lowering it by 7% or raising it by 8%. Similarly, changes in the return of AGZ have minimal impact on the nominal (Log-Stable Portfolio) return, lowering it by 5% or raising it by 9%. Government bonds (weighted at 10%) are less volatile. If the reader is familiar with asset class returns, this result is not surprising. Government bonds have much smaller returns (and less risk) than small cap stocks. However, the Tornado charts only look at the span between the 10^{th} and 90^{th} percentiles. In this sense, they fall short from inspecting the tails of distributions – which is where we want to look to see if risks, in the form of black swans, are hiding.

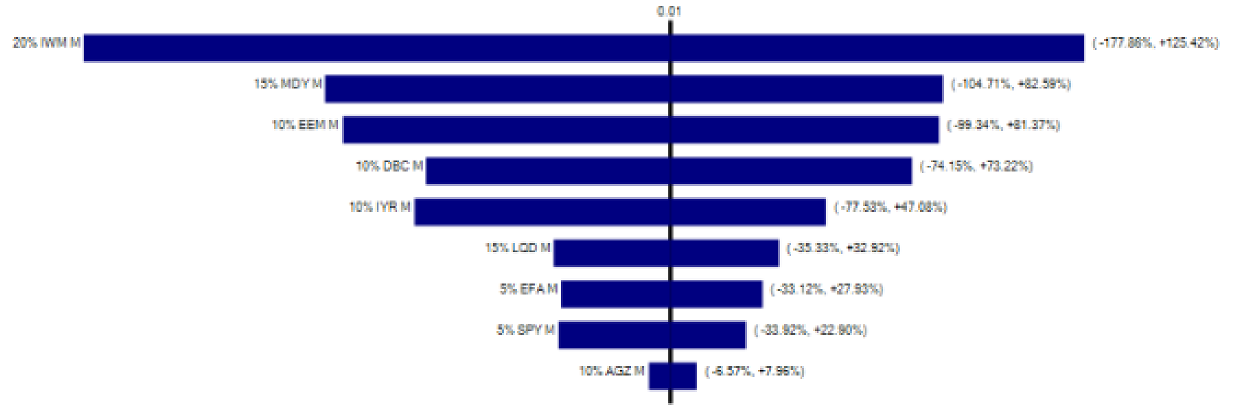

Figure 4A – Tornado Charts for Normal Portfolio

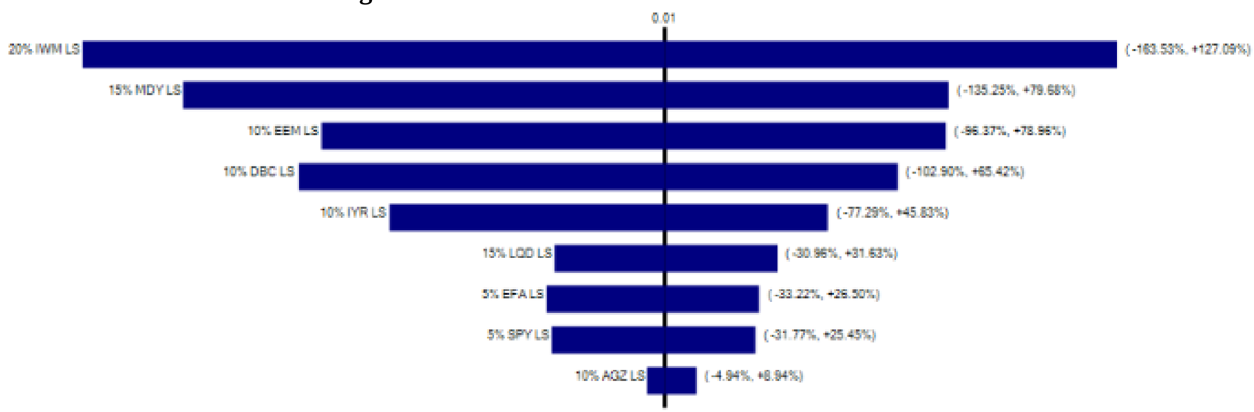

Figure 4B – Tornado Charts for Log-Stable Portfolio

Percent Point Functions (PPFs)

A PPF shows the probability of a random number being less than or equal to a particular point on the plot. For example in Figure 5 (left pane), there is a 50% probability the return will be 1% or less and there is a 90% probability the return will be 3% or less. The latter statement could be interpreted as a 10% probability the return will be greater than 3%.

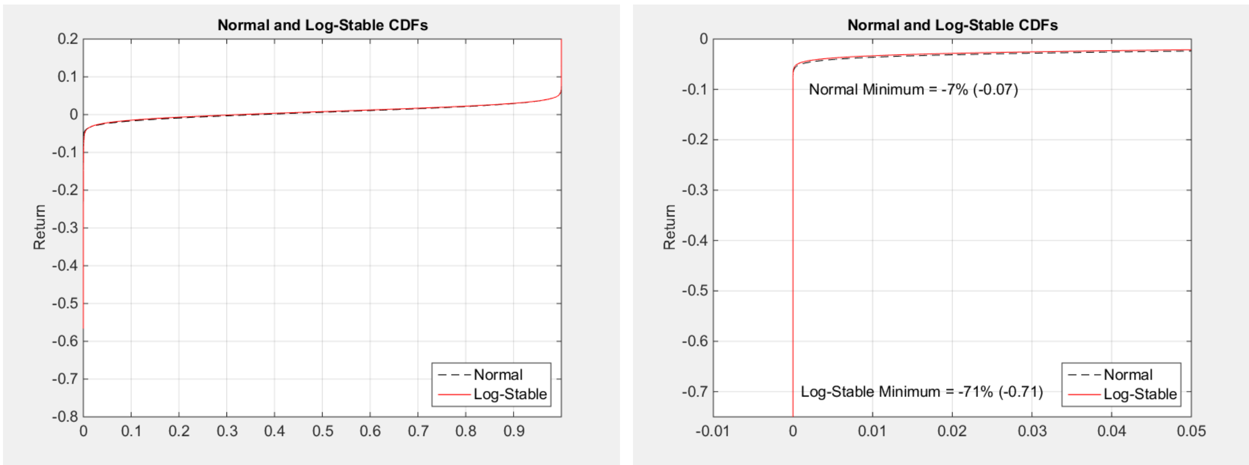

Figure 5 – PPFs for Normal and Log-Stable Portfolios

By examining Figure 5 (left pane), it appears as if the Normal and Log-Stable PPFs are identical. However, upon closer inspection (right pane), we see the Normal Portfolio tail stops near -7%, while the Log-Stable Portfolio tail continues down to -71%. Similarly, but not shown, the positive tails for the Normal and Log-Stable PPF returns are 7% and 16%, respectively. Once more, we've shown the log-stable distribution reveals larger risk, ten-fold for the portfolios. Therefore, the method of "PPF tail inspection" is a viable method to see if potential rare events (black swans) might be lurking in the data.

Rather than comparing Normal and Log-Stable Portfolios, we now blend two portfolios – one with historical monthly returns for all asset classes fit to normal distributions except for government bond returns (AGZ) which are fit with the log-stable distribution; the other portfolio will fit only real estate returns (IYR) to the log-stable distribution. These two (separate, but mixed) portfolios are selected knowing ahead of time government bonds have the lowest (historical) spread between risk and reward (-2% to 4%), while real estate has the highest (historical) spread (-31% to 29%). The reason for this is to examine the tails to see if any black swans might be identified, independent of risk and reward spread.

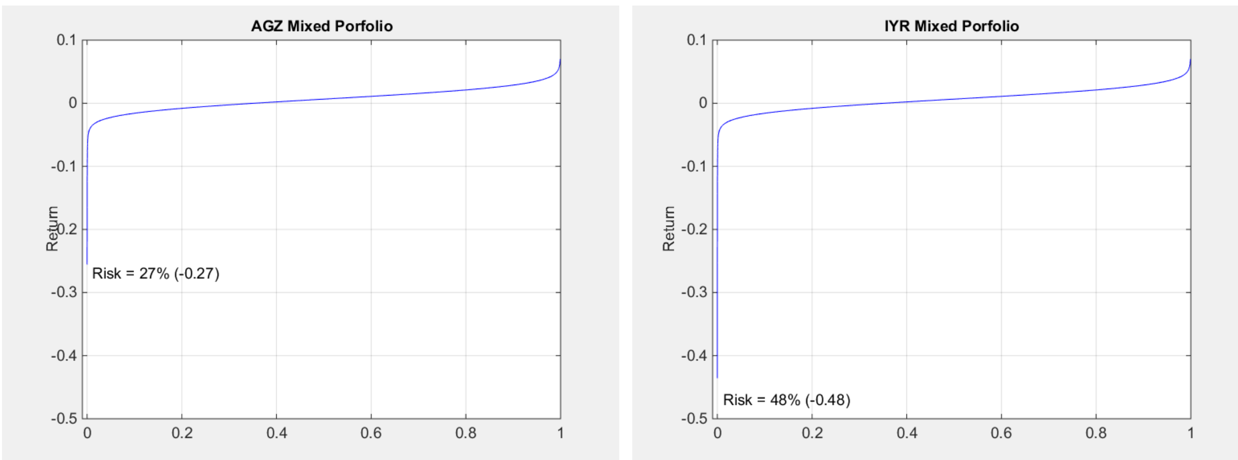

Figure 6 – PPFs for AGZ (left) and IYR (right) Portfolios

In the case of the AGZ Mixed Portfolio (Figure 6, left pane), we see it has significantly more risk (-27%) than the Normal Portfolio, which had 7% to the downside (Figure 3, left pane). Even by modeling a less volatile asset class with the log-stable distribution, we are able to see additional risk through the lens of the PPF tail. Of course, for the IYR Mixed Portfolio (Figure 6, right pane) the risk is more pronounced (-48%) due to its higher volatility and by virtue of being modeled with a log-stable distribution. The risk of IYR alone is driving the risk of the entire portfolio. In sum, the PPF tail exposes the possibility of the occurrence of a rare event.

Stochastic Optimization

Finally, we come to the method of stochastic optimization, where random variables appear in the formulation of the optimization problem thus producing a random objective function for which random iterates are employed to solve the problem. When optimizing (maximizing or minimizing) an objective function (portfolio), stochastic optimization assesses the range of each probabilistic input and selects whatever values are necessary to yield the desired result. For example, to find the minimum portfolio value, the minimum historical return of each asset class will be chosen.

For a small portfolio of only nine asset classes, it is rather simple to calculate deterministic minimum and maximum portfolio returns. Table 2 (below) shows minimum and maximum historical returns for each ETF. Based on the asset allocation we've been using (repeated in the table), the minimum and maximum portfolio returns are -22% and 14%, respectively.

Table 2 – Allocated Portfolio Minimum and Maximum Returns

ETF	Allocation	Minimum	Maximum
IWM	20%	-24%	14%
MDY	15%	-24%	14%
SPY	5%	-18%	10%
EFA	5%	-23%	12%
EEM	10%	-29%	16%
LQD	15%	-11%	13%
AGZ	10%	-2%	4%
IYR	10%	-38%	26%
DBC	10%	-29%	15%
	Portfolio	-22%	14%

The results of running stochastic optimization on this small portfolio were identical with the deterministic case, i.e. minimum and maximum returns of -22% and 14%, respectively. Using stochastic optimization for this sized problem is excessive. But, if a portfolio contains thousands of random inputs, including complex interconnections, it soon becomes intractable to perform these calculations with a spreadsheet, let alone by hand. The result of stochastic optimization is a list of all the inputs and the values that have been chosen so as to achieve either the minimum or maximum return.

The astute reader will wonder how the stochastic optimization risk and return range (-22% to 14%) relates to the prior results of the Normal (-7% to 7%) and Log-Stable (-71% to 16%) Portfolios. We've already discussed the normal distribution and how it naïvely characterizes risk. This explains why the risk and return range is lower than either of the other results. To explain the difference between the stochastic optimization and the Log-Stable Portfolio results, we recognize the historic lows and highs from Table 2 are bounded. For example, when stochastic optimization seeks a minimum, the algorithm selects minimal values of the inputs, which are the lower bounds. Likewise, for the maximum portfolio value, upper bounds are chosen. Stochastic optimization is dependent on the bounded values of the inputs. Compare this to the log-stable distribution which can return values beyond these limits, albeit with small (but non-zero) probability. Even though the historical data is bounded, the log-stable distribution fits the data with parameters which allows random numbers to be drawn in excess of these bounds – again, with small, but non-zero probability. Think of it this way, while the largest earthquake on record is magnitude 9.5, the possibility exists for a larger earthquake to occur – we just haven't experienced it, yet.

Recap

Tornado charts can show how varying an input impacts the output. This is useful to identify which inputs are influencing the output, but only between the 10^{th} and 90^{th} percentile. Examining the tails of the PPF shows how much impact rare events could have. Stochastic optimization displays the values of each of the inputs to achieve output extrema. Together, these three tools and their insights reduce the surprise of a black swan, rendering it gray.

TRAINING AND SIMULATION APPLICATIONS

Acquisition

The triple constraint in acquisition is cost, schedule, and performance. Any probabilistic inputs for these models should be fitted with fat tail log-stable distributions to help understand any exorbitant costs and extreme impacts of schedule slippage and poor performance.

Training Proficiency

Training proficiency is measured by its baseline effectiveness and the time and number of iterations invested in the exercise. Proficiency is augmented by media factors and instructional quality factors, but diminished by skill decay due to lack of training. Any or all of these stochastic inputs should be considered for being fitted with fat tail log-stable distributions.

Strategic Multi-Layer Assessment

Organizational models, social network models, time influence network models, information diffusion models, and text analysis models comprising strategic multi-layer assessment have numerous inputs, including event probabilities and event frequencies. Given the purpose of these models, one should most assuredly use fat tail log-stable distributions to expose any black swan events like the rise of Al Qaeda, Hazbollah, and most recently, ISIL.

SUMMARY

In this paper, we briefly discussed black swans and their attributes. More appropriately answering the first question from the "2015 I/ITSEC Black Swan Kickoff," we showed it's possible to identify potential black swans and in doing so, render them gray. Thus we can prepare, organize, train and equip for black swan resiliency.

Furthermore, we showed the log-stable distribution is preferred to the normal distribution when it comes to modeling data that includes rare events lying far from the mean. The log-stable achieves this by assigning a non-zero probability of occurrence with its fat tail, whereas the normal distribution assigns a negligible probability due to its thin tail. We showed practical (financial) applications of log-stable modeling for both individual data sets (S&P 500) as well as a portfolio comprised of data sets (ETFs).

Finally, we discussed how three tools (Tornado charts, PPFs, and stochastic unconstrained optimization) and their insights can reduce the surprise of a black swan, rendering it gray.

In the end, we have shown how modeling and simulation can be used to analyze and prepare or create a black swan and in practice, we've developed models and simulation tools that enable the analysis.

ACKNOWLEDGEMENTS

We wish to thank Dr. Paul Kaplan (Morningstar) and Dr. John Nolan (University of Virginia) for personal email correspondence. We wish to recognize Investools / TD Ameritrade as the database from which we were able to obtain historical monthly data for each asset class.

REFERENCES

Fama, E. (1963). Mandelbrot and the Stable Paretian Hypothesis. Journal of Business, Vol. 36, No. 4, 420-429.

Kaplan, P. (2012). Frontiers of Modern Asset Allocation. Hoboken, NJ: John Wiley & Sons, Inc.

Kaplan, P. (2008). Using Fat Tails to Model Gray Swans. Retrieved from http://morningstardirect.morningstar.com/clientcomm/LogStableDistributions.pdf

Mandelbrot, B. (1963). The Variation of Certain Speculative Prices. Journal of Business, Vol. 36, No. 4, 394-419.

Markowitz, H. (1952). Portfolio Selection. Journal of Finance, Vol. 7, No. 1, 77-91.

Markowitz, H. (1979). Approximating Expected Utility by a Function of Mean and Variance. The Economic Review, Vol. 69, No. 3, 308-317.

Markowitz, H. (1999). The Early History of Portfolio Theory: 1600–1960. Financial Analysts Journal, Vol. 55, No. 4, 5-16.

Nolan, J. (2009b). User Manual for STABLE 5.1: Matlab Version. Retrieved from www.robustanalysis.com/MatlabUserManual.pdf.
Savage, S. (2009). The Flaw of Averages. Hoboken, NJ: John Wiley & Sons, Inc.
Taleb, N. (2010). The Black Swan – The Impact of the Highly Impossible. NY: Random House Inc.
Veillette, M. (2015). STBL: Alpha stable distributions for MATLAB. Retrieved from http://www.mathworks.com/matlabcentral/fileexchange/37514-stbl--alpha-stable-distributions-for-matlab/all_files
Wright, C. (2007, March/April). Tail Tales. Chartered Financial Analyst Institute Magazine.

APPENDIX – MATHEMATICAL PROPERTIES OF THE LOG-STABLE DISTRIBUTION

The log-stable distribution is frequently used to model investment returns. Returns are expressed in decimal form, where negative returns represent losses and positive returns represent profit. We then normalize the returns by adding one and taking the natural log of the result. Once in this form, the returns conform to a stable distribution. The probability density function (pdf) for a (fat-tail) stable distribution is

$$f\left(x:\alpha,\beta,\gamma,\delta\right)=\frac{1}{\gamma}g\left(\frac{x-\delta}{\gamma};\alpha,\beta\right)$$

where

α represents the "fatness" of the tails and is in the range between 0 and 2, with 2 being a normal distribution. Also, if alpha is less than 1, then the mean of distribution is infinite.
β represents the skewness of the distribution and lies within the range of -1 to 1, where -1 signifies fully left-skewed and +1 signifies fully right-skewed. If beta is 0, the distribution is symmetric.
γ represents the scale of the distribution and is positive. If alpha = 2 (normal), then gamma squared is one-half the variance.
δ represents the location of the distribution. If alpha > 1, then delta is the mean of distribution.