Bamboozled by Mathematical Models

By John F. McGowan

Version: 1.1.3
Start Date: February 13, 2009
Last Updated: February 16, 2009
Home URL: http://www.jmcgowan.com/bamboozled.pdf

Mathematical models have solved important problems and hold the promise of solving major outstanding problems such as the need for cheap, plentiful energy. They can also be used to intimidate and bamboozle the unwary as the current financial crisis illustrates.

Introduction

A mathematical model uses mathematical language to describe a system. A very simple mathematical model is the mathematical rule for the speed of an object dropped near the Earth's surface:

$$V = gt$$

Where the symbol V represents the speed of the falling object usually measured in meters per second or feet per second, the symbol g is 9.8 meters per second per second (or 32 feet per second per second in English units), and the symbol t is the time since the object was dropped. What this mathematical model means is that one second after the object is dropped ($t=1.0$) the object will be falling at 9.8 meters per second (32 feet per second). Two seconds after the object is dropped, the object will be falling at 19.6 meters per second (64 feet per second in English units). This mathematical model, one of the simplest, is usually attributed to Galileo. This model works fairly well although in reality one must include the effect of air resistance to make an exact prediction of the speed of the falling object. If the object is massive and dense the effect of air resistance can be quite small. This mathematical model can be plotted on a graph:

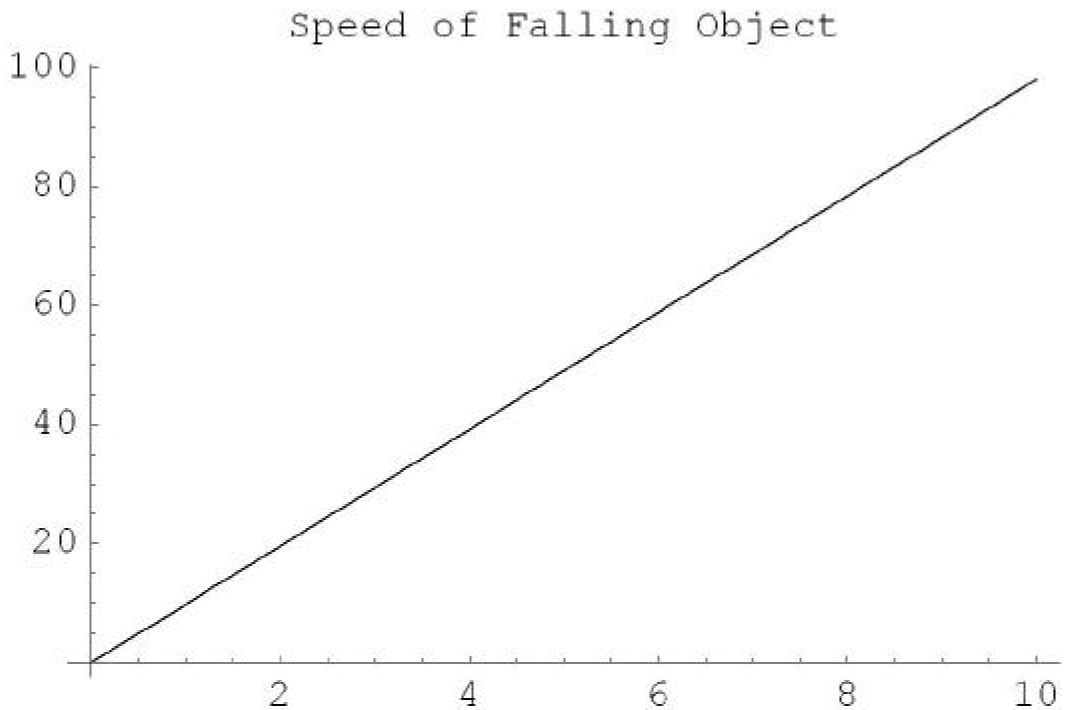

Figure 1 Speed of a Falling Object

Mathematical models can be more complex. A slightly more complex mathematical model is the formula for the distance traversed by an object dropped near the Earth's surface:

$$D = \frac{1}{2} g t^2$$

Where the symbol D represents the distance that the object falls, the symbol g is 9.8 meters per second per second (32 feet per second per second in English units), and the symbol t represents the time since the object was dropped.

This mathematical model can also be plotted on a graph:

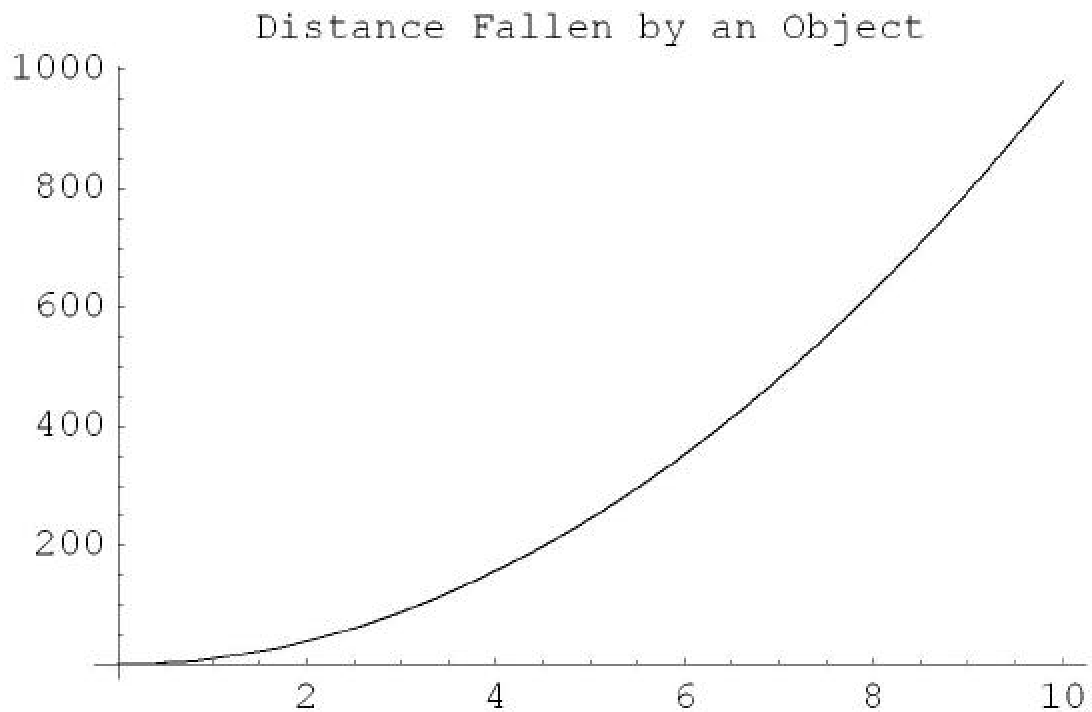

Figure 2 Distance Fallen by an Object

Mathematical models can be even more complex. A more complex mathematical model is the Gaussian or Normal function, also known as the Bell Curve. The mathematical formula for the Gaussian is:

$$G(x, \sigma) = \frac{e^{-\frac{x^2}{\sigma^2}}}{\sqrt{2\pi}\,\sigma}$$

Where the symbol *e* is the special mathematical constant 2.71828, the symbol π (the Greek letter PI) is 3.14159 (the ratio of the circumference to the diameter of a circle), *x* is the independent variable, and σ (the Greek letter SIGMA) is the *standard deviation* of the Gaussian. The standard deviation is a measure of the spread or width of the Gaussian peak.

This cryptic mathematical formula with Greek letters and mathematical symbols can be intimidating and confusing. The mathematical formulae for more complex models can be even more cryptic, intimidating, and confusing.

The Gaussian is widely used, overused in fact. The Gaussian with a standard deviation of 1.0 can be plotted on a graph as follows:

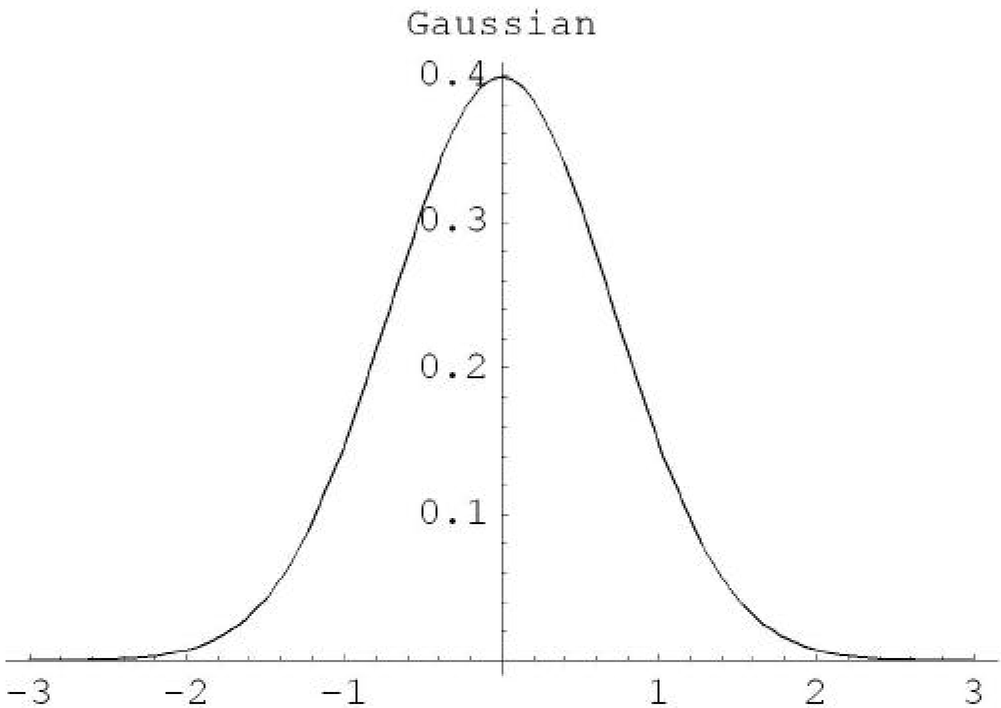

Figure 3 The Gaussian Function

Notice that the Gaussian falls off very rapidly as the location moves away from the central peak.

The Gaussian occurs in the real world, for example in flipping a coin. The Gaussian predicts the probability of the number of heads is a large number of coin tosses. Take a coin and flip the coin many times, for example one hundred (100) times. Count the number of heads that occur (for example, fifty-eight). Subtract fifty (100 times the one-half chance of a head on a single coin toss) from the number of heads (for example, 8 = 58 – 50). Divide this number by the square root of the number of coin tosses (one hundred in the example) times the probability of a single head (½) times the probability of a single tail (½). This is the square root of twenty-five (25) or five (5). In the example, this number is (58 – 50)/5 = 1.6. The Gaussian is approximately the probability of this number (1.6 would be *x*); the probability of fifty-eight heads in one hundred coin tosses is 0.03 from the Gaussian. In the example, the probability is about *G(1.6,1.0)* which is about 0.03 or three percent.

The Central Limit Theorem, a famous mathematical theorem, says that the average of a random variable over many samples approaches a Gaussian distribution as the number of samples tends to infinity. The coin toss example is one case of the Central Limit Theorem. Partly because of the Central Limit Theorem, the Gaussian is often used in mathematical models.

Mathematical models can be extremely complex. Extremely complex mathematical models today are often sums and products or other simple combinations of large numbers of simpler mathematical functions such as the Gaussian. The Gaussian is used heavily in the textbook Hidden Markov Model (HMM) speech recognition algorithms.

The mathematical models in the speech recognition now used for many telephone help lines have thousands of terms and thousands of parameters. The constant g in the simple examples above is a single parameter that is experimentally determined. There is much room for improvement in the mathematical models used in speech recognition today. The compressed video in a DVD video disc or a BluRay video disc can be considered a series of parameters for a complex mathematical model of the video that is nonetheless much simpler than the uncompressed video. The mathematical models used in video compression work extremely well and may be nearing their theoretical limits although significant improvements in the perceived quality may still be possible.

Examples such as the laws of gravity, digital video compression, and speech recognition show that mathematical models can work and solve (or partially solve) real problems. Mathematical models can also be wrong. They can be used to bamboozle or intimidate the unwary.

Bamboozled

When I was in graduate school, I received a phone call out of the blue from a headhunter looking for physicists and similar people for a job somewhat mysteriously called a “quantitative research analyst” on Wall Street apparently applying advanced mathematical methods to high finance. I was rather intrigued especially given the rather high salary that was mentioned, considerably more than I was making at the time as a graduate student. Having been following the rather unsettling saga of Drexel Burnham Lambert for several years as a sort of hobby, I rather innocently asked “well, is investment banking cutthroat?”. “No, of course not,” said the headhunter. I was somewhat skeptical even at that point.

Nonetheless intrigued both by the technical problem and the potential salary, I investigated further. I read several books and traveled to New York City to talk with a few headhunters while visiting the East Coast[1,2,3,4]. One conversation in particular stuck in my mind. One headhunter described how he had placed a recent Ph.D. as a "quantitative research analyst" (now better known as a quant) with some Wall Street firm. This person lasted six months and quit. The headhunter asked this person why he quit. He said "Well, it is bogus research." I heard things like this from some other sources as well. Now, this is what is called "anecdotal evidence" and scientists typically frown on it.

Somewhat skeptical of the headhunter, I tracked down a relative who worked in finance who assured me that "investment banking is *very* cutthroat". This news was perhaps not a total surprise.

I decided that I would almost certainly not fit in with the culture of Wall Street and pursued a career in mathematically oriented computer software such as video compression. Nonetheless, I kept an eye on the murky and high-paying world of quantitative finance. Over the years many physicists have trekked from the sheltered groves of academia (well, not so sheltered in the cutthroat world of particle physics, but that is another story) to Wall Street[5].

In 1998, quantitative finance suffered a slight black eye when the Long-Term Capital Markets hedge fund which included several luminaries of the quantitative finance world (Nobel Laureates Robert Merton and Myron Scholes of the Black-Scholes option model) failed[6]. Allegedly, Long-Term almost crashed the entire financial system, almost taking several major Wall Street investment banks with it. Fortunately, so they say, the Federal Reserve stepped in to rescue the Wall Street folks, mostly the same firms involved in the current mortgage-backed securities meltdown.

Long-Term Capital Markets allegedly relied upon sophisticated mathematical models that worked great for several years and then failed miserably. One might think this would lead Wall Street, not to mention regulators like the Federal Reserve, to question the use of such models. Apparently not. Similar models, known generically as Value at Risk (VaR) models, are playing a similar role in the current financial crisis, which has the dubious distinction of replaying the Long-Term Capital Markets story on an even larger scale. Once again the

Federal Reserve and the US Treasury are riding to the rescue, with rather less success this time.

Former quantitative trader Nassim Nicholas Taleb has produced a barrage of articles, web sites, books, and outraged appearances on video denouncing the VaR models for years. He apparently enjoyed a huge audience at the World Economic Forum in Davos recently. He asserts that the VaR models or at least many of them assume a Gaussian distribution (or something similar) but the real distributions have long non-Gaussian tails. Every once in a while there is an outlier like the housing bubble popping that wrecks the neat Gaussian model[7,8].

Here is a relevant quote from *When Genius Failed: The Rise and Fall of Long-Term Capital Management* by Roger Lowenstein (published in 2000, page 72):

[Eugene] Fama's thesis, which he undoubtedly discussed with [Myron] Scholes, was rife with implications for the future of Long-Term Capital. In contrast to the idealized markets in the models, Fama warned, real-life ones experienced "discontinuous" price changes (those nasty jumps) and a higher probability of large losses; indeed, "such a market is inherently more risky"

By the time Long-Term was formed, it was well documented that all *financial assets behaved like the stocks that Fama had studied.* ***Mortgage securities [emphasis added]*** *might usually behave as the model predicted, but there would come a day when – with no warning at all – they would leap off the charts. As Fama put it, "Life always has a fat tail."*[9]

In graduate school, I spent about two years trying to construct a mathematical model of the sub-atomic process in which an electron and positron (anti-electron) annihilate to produce four pions. The pion is the subatomic particle that is thought to be the primary carrier of the strong nuclear force, the force that powers the Sun, thermonuclear bombs, and some radioactive decays. The goal was to measure the rho form factor, which is a measure of the shape of the rho meson, which is also thought to be a carrier of the strong nuclear force. I was unsuccessful. The mathematical models never agreed with the data. I did find a way to put an upper limit on the rho form factor that did not rely on mathematical models, but an upper limit is not as satisfying as a measurement.

In the process, I learned many of the perils and pitfalls of mathematical models, one of which is that real world distributions are often not Gaussian even though mathematical modelers often blindly assume a Gaussian distribution based on a naïve application of the Central Limit Theorem. The Gaussian drops off very rapidly, exponentially, as one moves away from the central peak. The probability of an event even ten standard deviations from the peak is essentially zero. In the real world, one often encounters outliers far from the peak. The famous Black-Scholes option pricing model, for example, assumes a Gaussian distribution for the "returns" of the stock or other assets that the option applies to. One may perhaps see why even in graduate school I had some skepticism about quantitative finance.

This is the plot for the function

$$\frac{1.}{x^2 + 1.}$$

which has long tails:

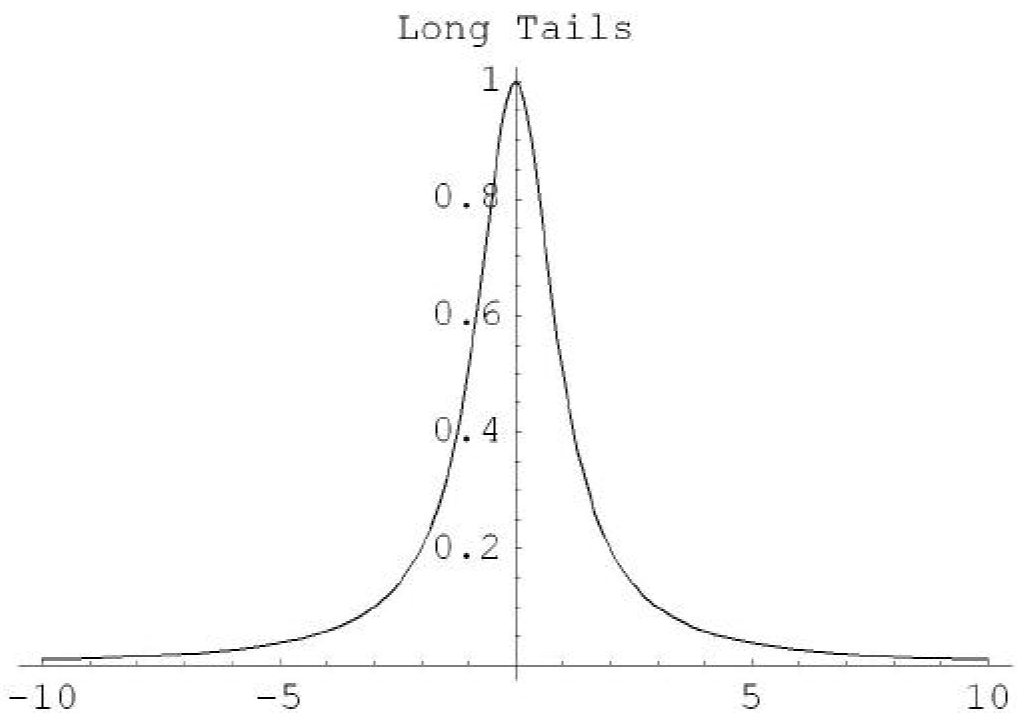

Figure 4 A Function with Long Tails

Outliers or long tails are not unique to finance. In fact, it is common in many fields to encounter outliers or non-Gaussian tails. Even where a distribution is mostly Gaussian, occasional non-Gaussian outliers often occur. There are many causes for these anomalies: measurement errors, mistakes such as transpositions of digits during hand entry or copying of quantitative data, rare physical processes that are not understood, and so forth. In practice, outliers are often removed from the data particularly when they can be identified as measurement errors. There is an entire field of *robust statistics* devoted to statistical methods such as the median that are robust against outliers or other differences between the actual distribution of the data and an idealized mathematical model such as the Gaussian[10].

Long tails is actually only one of many problems with the mathematical models used in finance and economics.

Conclusion

Mathematical models can work. When you struggle with a telephone help line that uses computerized speech recognition, you are talking to a mathematical model that needs improvement. Whenever you watch a DVD, a BluRay disc, or a YouTube video, you are watching a kind of mathematical model that works very well. Mathematical models of electromagnetic and nuclear forces may enable us to design fusion reactors, possibly even the size of a basketball and costing a few thousand dollars that can provide cheap and plentiful energy. Unfortunately there are many perils and pitfalls to mathematical models. Some mathematical models can be wrong or harbor very serious flaws as the current financial crisis shows.

The great tragedy of the current financial crisis is that so much time and effort, money and talent have been spent on mathematical models that not only didn't work but have proven very harmful. Instead of solving problems like nuclear fusion or curing cancer, mathematical models have been put to work at best as the drivers of a fiasco and at worst as a smokescreen for a massive fraud that may drag the world into war as desperate and frightened people fight over dwindling resources.

The failure of financial models in the current mortgage backed securities crash, Long-Term Capital Markets, and a number of other derivative securities fiascoes illustrates the danger of relying on gold-plated resumes, credentials, impressive academic records, and even Nobel prizes. Both the data compression and speech recognition fields also feature major failures and scandals such as Lernout and Hauspie that involved people and organizations with impressive academic and professional credentials. It is essential to evaluate the ideas, the actual mathematical models, and the actual software or hardware where a working prototype exists or is claimed to exist.

Computers now have enormous unused power. It is hard even to come up with something to use the full power of a Gigahertz CPU. Mathematical models for speech recognition and other unsolved problems can put that enormous power to beneficial use. To achieve this goal, people need to appreciate the perils and pitfalls of mathematical models, and avoid being intimidated or bamboozled by seemingly sophisticated models that either don't work or hide fatal flaws.

About the Author

John F. McGowan, Ph.D. is a software developer, research scientist, and consultant. He works primarily in the area of complex algorithms that embody advanced mathematical and logical concepts, including speech recognition and video compression technologies. He has many years of experience developing software in Visual Basic, C++, and many other programming languages and environments. He has a Ph.D. in Physics from the University of Illinois at Urbana- Champaign and a B.S. in Physics from the California Institute of Technology (Caltech). He can be reached at jmcgowan11@earthlink.net.

---

[1] John Hull, *Options, Futures, and Other Derivative Securities*, Prentice-Hall, Englewood Cliffs, New Jersey, 1989
[2] Richard M. Bookstaber, *Option Pricing and Investment Strategies*, 3rd Edition, Probus Publishing Company, Chicago, IL, 1991
[3] Rajna Gibson, *Option Valuation: Analyzing and Pricing Standardized Option Contracts*, McGraw-Hill, New York, 1991
[4] Peter L. Bernstein, *Capital Ideas: The Improbable Origins of Modern Wall Street*, The Free Press, A Division of MacMillan Inc., New York, 1992
[5] Emanuel Derman, *My Life as a Quant: Reflections on Physics and Finance*, John Wiley and Sons, Hoboken, New Jersey, 2004
[6] Roger Lowenstein, *When Genius Failed: The Rise and Fall of Long-Term Capital Management*, Random House, New York, 2000
[7] Nassim Nicholas Taleb and Pablo Triana ,"Bystanders to this financial crime were many", Financial Times (FT.COM), Published: December 7 2008 19:18 | Last updated: December 7 2008 19:18 (URL: http://www.fooledbyrandomness.com/Ft-Bystanders.pdf, Accessed February 14, 2009)
[8] Nassim Taleb, *The Black Swan: The Impact of the Highly Improbable,* Random House, New York, 2007
[9] Roger Lowenstein, *When Genius Failed: The Rise and Fall of Long-Term Capital Management*, Random House, New York, 2000, page 72
[10] Peter J. Huber and Elvezio M. Ronchetti, *Robust Statistics,* John Wiley and Sons, 2009