Why Fat Tails Matter - A Primer

Introduction

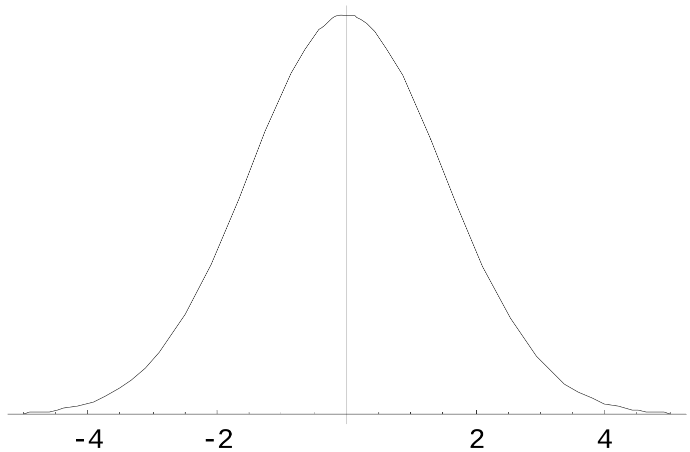
There is a joke about State Lotteries that says: "The Lottery is a TAX on people who are bad at math". Sadly, this is true. It is also a regressive tax because it takes a bigger portion of the income of poor people who typically are the least educated and therefore the least able to determine just how bad a financial opportunity the lottery is.

The math is really not all that difficult. It requires the understanding of a simple concept: The expectation of an uncertain outcome. Here is a simple example. Suppose you find yourself in a situation where there are three possible outcomes, A, B, and C, each with a different reward. Suppose I, being a benevolent and omniscient dictator, know and tell you in advance the exact probability of each outcome and I give you the opportunity to repeat your choice many, many times. If you take each reward and multiply it by each probability then add up the results you have what is known as the mathematical expectation. Here is an example: The payoffs are $10, $20, and $50; the probabilities are, respectively, 60%, 30% and 10%. The expectation is $6 + $6 + $5 = $17.

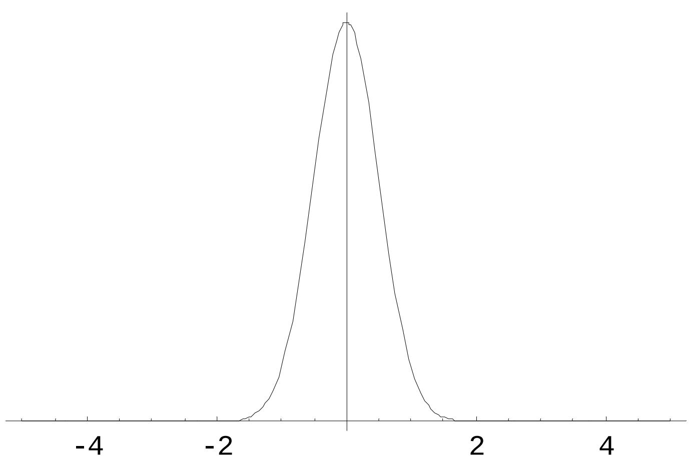
Now suppose I charge you a fee to play this game. How much, given that you can play it repeatedly as often as you like while the rules remain the same, are you willing to pay? This little thought experiment is at the heart of the lottery joke. Ignoring the blizzard of possible ways they allow you to win, let's just take the one that gets headlines. Suppose the jackpot is $10,000,000. The lottery people announce the number of tickets sold and the odds of winning. Suppose the odds are one in 24,000,000, a decimal fraction of .00000004166667. Multiply the payoff times the probability and you have just less than 42 cents. What is the price of a lottery ticket?

The simple conclusion is that people who buy lottery tickets are overpaying for the opportunity. This is the result of failing to understand the notion of an expectation. There is even a germ of common sense in it: If someone offers you the chance to win a dollar based on the flip of a fair coin, DON'T PAY MORE THAN 50 CENTS FOR THE CHANCE!!!

Of course, life is more complex than flipping coins and omniscient dictators are hard to find, but the basic reasoning does not change. Many things in life are examined statistically and for the past 350 years a very useful and powerful tool has been used for such inquiries. It is called a probability distribution function (pdf) and it looks like this:



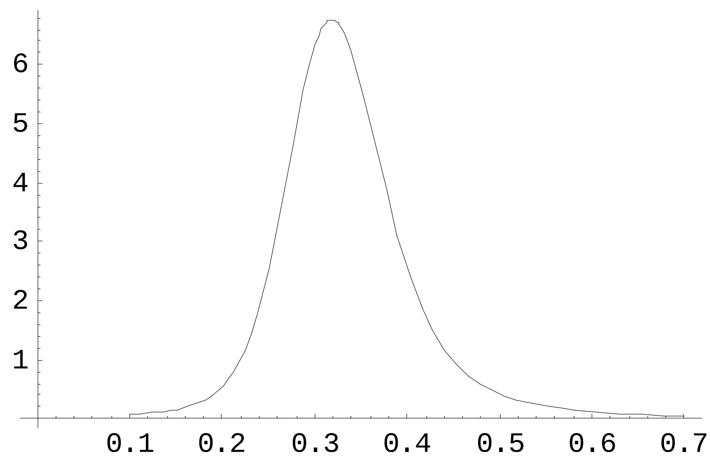
All this does is provide a graphical representation of all the outcomes that can happen, arranged in order of the product of their probabilities times their payoffs. The peak occurs where the most likely is located, which on the x-axis is the expectation. For this illustration the expectation is zero. There is a dispersion of possibilities around the expectation that accounts for the area under the curve that is away from the center. Below is a case with the same expectation but less dispersion, also known as variation. Literally and in formal statistical meaning the area under the curve represents the variance from the expectation. This may be viewed as how likely or unlikely it is that our expectations will be met.



These two curves have been created using the mathematics of the famous "bell curve", also known as the normal distribution. As important and powerful as this technology is, it has limitations. Several technical matters that we won't consider here constrain the use of it. First, it is symmetrical, meaning that there must be an equal number of outcomes on either side of the mean (another term for the expectation). Second, it rapidly descends to the x-axis creating something that statisticians refer to as a particular kind of "asymptotic behavior". This rapid decline precludes finding or considering a large number of outcomes of any size or a small number of large outcomes that are located far from the expectation. Such things are known as "outliers". Another way of saying this is that the normal distribution is "compact" around the mean. All observations must be within a certain range of the mean to permit the normal distribution to apply accurately.

Nature, while often exhibiting normal behavior, also shows us many other cases where outliers are persistent, meaningful and may dominate the probability calculation. Until very recently, the computation of probabilities for non-normal distributions has been difficult or impossible. This difficulty has been recently overcome and the study of outliers or conditions in which outliers play a significant role is now within reach.

The presence of outliers creates a condition known as "Heavy Tails" or "Fat Tails". This means that, because there are observations far from the mean and/or those observations are perhaps individually large, the tails DO NOT rapidly descend to the x-axis, rather they extend some distance from the expectation before approaching the x-axis as in the example below.

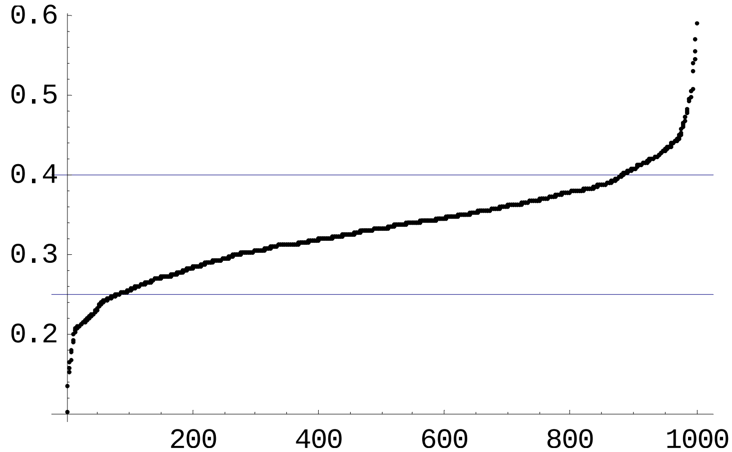


The stable approach to data

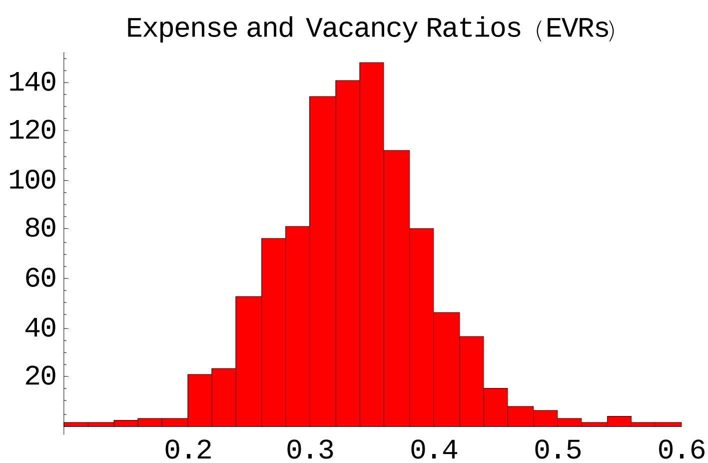
We will explore and illustrate this with a particular set of data. Here are 1000 observations of reported expense/vacancy ratios (EVRs) for apartment buildings in San Francisco. The first few look like this.

```
0.31745
0.344479
0.334084
0.321322
0.274807
0.322092
```

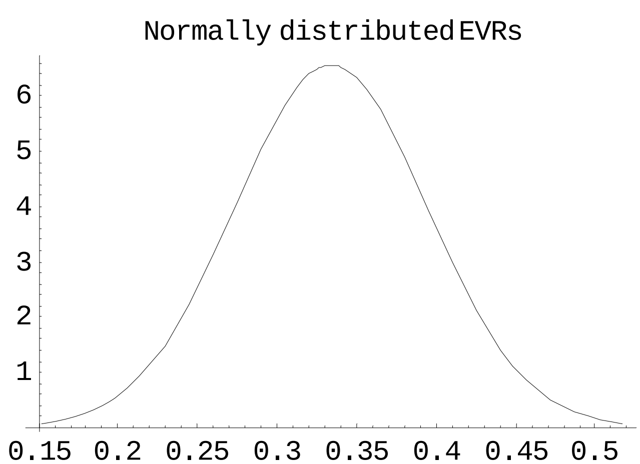
It is useful to look at the range of the observations and plot them. We see that most of them range between .25 and .4



When we view the data in histogram form we note that (a) the distribution of expense ratios is not symmetric and (b) that the distribution has a long right tail.



A special and useful property of the normal distribution is that you can create a pdf for it by knowing only the mean (expectation) and variance, known as the first two "moments" of the distribution. Assuming (naively) for now that the expense ratio observations *are* distributed normally, we can create and plot the pdf for such a distribution from its first two moments.



A companion and equally important tool is a Cumulative Distribution Function (CDF) that permits us to calculate the portion of the distribution that is contained under just a fraction of the pdf. The practical meaning of this is to permit us to compute probabilities of particular events.

With the above, continuing to believe that our distribution is normal, we can, using the CDF, assess the probability that any specific reported expense ratio will occur. Some examples of this are shown below.

evr	P(evr)
0.25	0.0839818
0.3	0.288662
0.35	0.604179
0.4	0.861185
0.45	0.971748
0.5	0.99682

This says that 99.7% of our EVR observations are at or below 50% while only 08.4% of them are at or below 25%. Thus, about 91% of them are between 25% and 50%. If we are given a capitalization rate that is based on an evr of 30% we see from the table above that only 28.9% of the buildings have an evr that low or lower. From this we can make a subjective assessment of the reliability of the capitalization rate. But how reliable is the model we employed to make this claim?

Two new terms are needed at this point. "Skewness" describes the extent to which the distribution is non symmetrical. "Kurtosis" measures the fatness of the tails. A check of the EVR distribution shows it is not normal because normal distributions are symmetrical (skewness = 0) and have "skinny" tails (kurtosis excess = 0). Neither is the case for this data.

Skewness	0.221308
Kurtosis Excess	1.07313

The stable approach to data

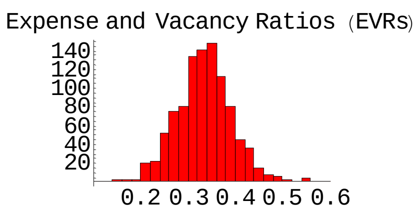
The assumption of normality imposes a set of strong conditions on the data. Included in these are symmetry, thin tails and a finite variance. Relaxing the normality assumption permits a better view of the data and, hopefully, the world from which it is drawn. The normal distribution is a special case of the family of Stable-Paretian (SP) distributions discussed at length elsewhere. SP distributions are characterized by four (4) parameters. One of these is α, the index of stability, which provides a measure of tail behavior. Another is β, a skewness parameter. For the (special) normal case $\alpha = 2$ and $\beta = 0$.

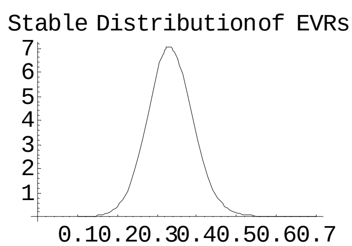
Because of recent technological developments, we can now estimate Stable parameters. For our data we see that α is below 2 and β is nearly 1, meaning that it has a heavy right tail. This means that there is more probability in the right (higher in this case) end of the distribution.

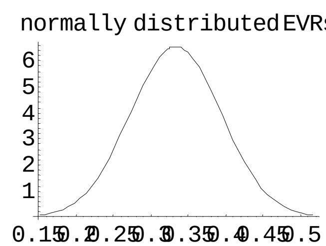
$\alpha = 1.87561$ $\beta = 0.36048$ $\gamma = 0.0400139$ $\delta = 0.334212$

Let's change tactics and deal with dataset under the assumption that it has a Stable distribution that is not normal.

Notice the similarity in the shape of the histogram of the actual data and the shape of a plot of the *Stable* pdf. Contrast that with the assumption of normality made earlier as shown in the far right plot below. The assumption of normality distorts our view of the data away from its actual shape. Note below how much better the SP distribution matches the histogram than the normal distribution does.







Comparing expense ratio probabilities provides the following differences in our estimation of how likely a building expense ratio is. In this case the normal assumption was not far from the stable estimates.

evr	Normal Prob	Stable Prob	Difference
0.25	0.0839818	0.0731309	0.0108509
0.3	0.288662	0.281732	0.00692929
0.35	0.604179	0.620898	-0.0167186
0.4	0.861185	0.875039	-0.0138543
0.45	0.971748	0.969291	0.00245782
0.5	0.99682	0.990765	0.00605495

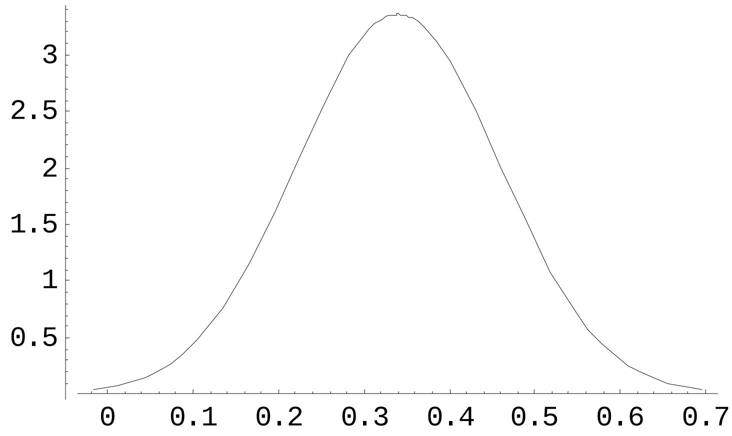
Recall the estimates stable parameters included an α that was close to the normal. This is why our normal assumption produced estimated probabilities reasonably close to those calculated with parameters from our stable fit.

EVR stable fit = {1.87561, 0.36048, 0.0400139, 0.334212}

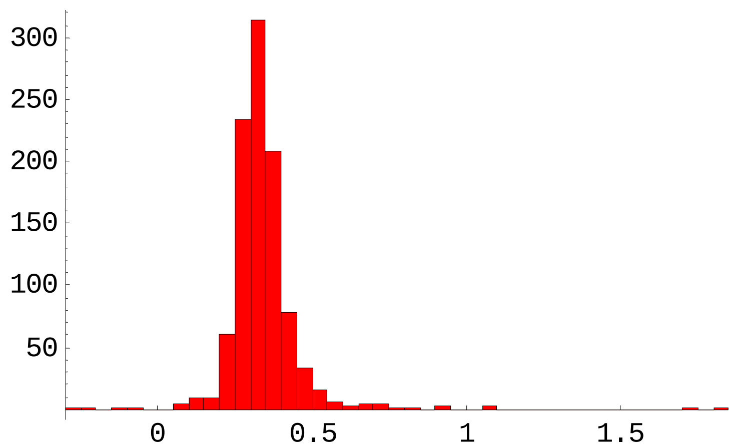
But let's assume we encounter a market in which α is considerably below 2. We now create a dataset for such a market using a random number generator. We assume this market has an α = 1.5

random number stable fit = {1.5364,0.356897,0.0400721,0.331344}

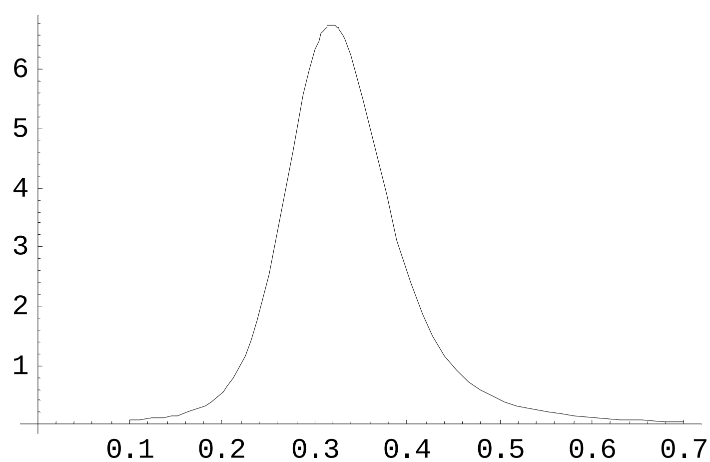
Applying our convenient normality assumption produces the familiar bell-shaped plot



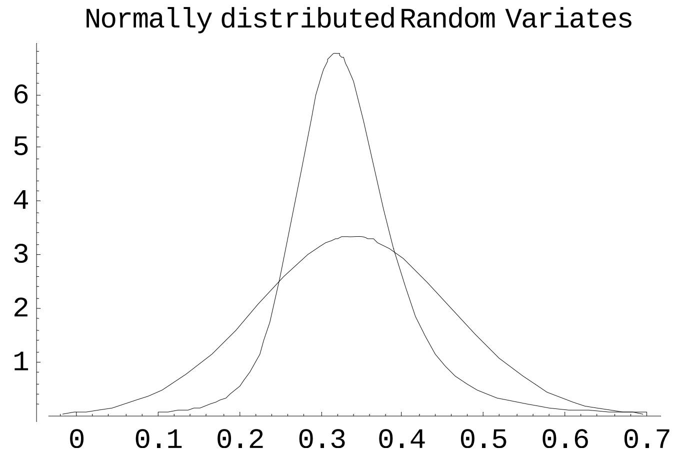
But the histogram indicates that this data is shaped rather differently.

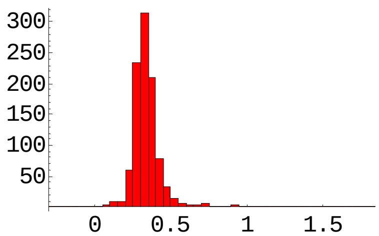


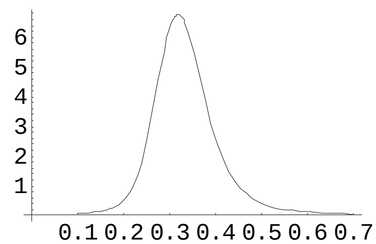
And a plot of the stable pdf looks more like the histogram than the normal pdf does. The reason for this is the superior fit of the stable pdf that considers the skewness and kurtosis that normal excludes.

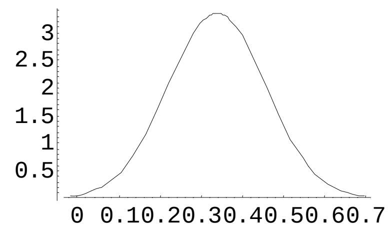


Placing them on the same plot or side-by-side is sobering









Calculating the probabilities both ways and computing the difference shows the magnitude of the error of the normal assumption.

evr	Normal Prob	Stable Prob	Difference
0.25	0.224803	0.093159	0.131644
0.3	0.36868	0.318036	0.0506444
0.35	0.534016	0.641199	- 0.107183
0.4	0.693603	0.843466	- 0.149864
0.45	0.822987	0.925419	- 0.102432
0.5	0.911095	0.957677	- 0.0465825

Returning to our original metaphor of a game with known probabilities, suppose it became known that I, as benevolent and omniscient dictator (or just your financial advisor), provided you with probabilities that I calculated based on the assumption of normality. Suppose that you accepted those numbers and on that basis decided the amount you were willing to pay to enter the game. Then suppose that someone else came along with the ability to estimate parameters under a fat tailed, stable assumption? The table above should show you that the price you should be willing to pay is very different.

Stretching our lottery metaphor to the limit, we might conclude that the assumption of normality is a tax on people who can't estimate stable parameters.