

Distributional Entropy, Information, and Fat-Tails

Navanit Arakeri

arakeri@stanfordalumni.org

Draft. April 2008

Introduction

This essay is in three parts. In the first section I show that fat-tailed distributions have a higher entropy compared to thin-tailed ones. In the second section I model capital markets as repositories of information and introduce the notion of entropy as a measure of information content. I then show that an increase in information will lead to an increase in kurtosis. In the third and final section I list some testable predictions based on this hypothesis.

1. Distributional Entropy and Kurtosis

I use the Student's-T distribution and vary the parameter 'v' to shift between thin and fat tails. The Student's-T distribution has a probability density function given by:

$$\frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\nu\pi} \Gamma(\frac{\nu}{2})} \left(1 + \frac{x^2}{\nu}\right)^{-\frac{(\nu+1)}{2}}$$

With Kurtosis:

$$\frac{6}{\nu - 4} \text{ for } \nu > 4$$

and Entropy:

$$\frac{\nu+1}{2} \left[\psi\left(\frac{1+\nu}{2}\right) - \psi\left(\frac{\nu}{2}\right) \right] \\ + \log \left[\sqrt{\nu} B\left(\frac{\nu}{2}, \frac{1}{2}\right) \right]$$

with ψ being the DiGamma function and B being the Beta function.

I vary the tail exponent parameter 'v', to move from a benign, near-gaussian distribution to a more vicious near-cauchy distribution with fat-tails. I compare the behavior of the entropy and excess kurtosis.

Figure 1 below shows the relationship computed numerically, with a Taylor-series expansion used to compute the DiGamma function:

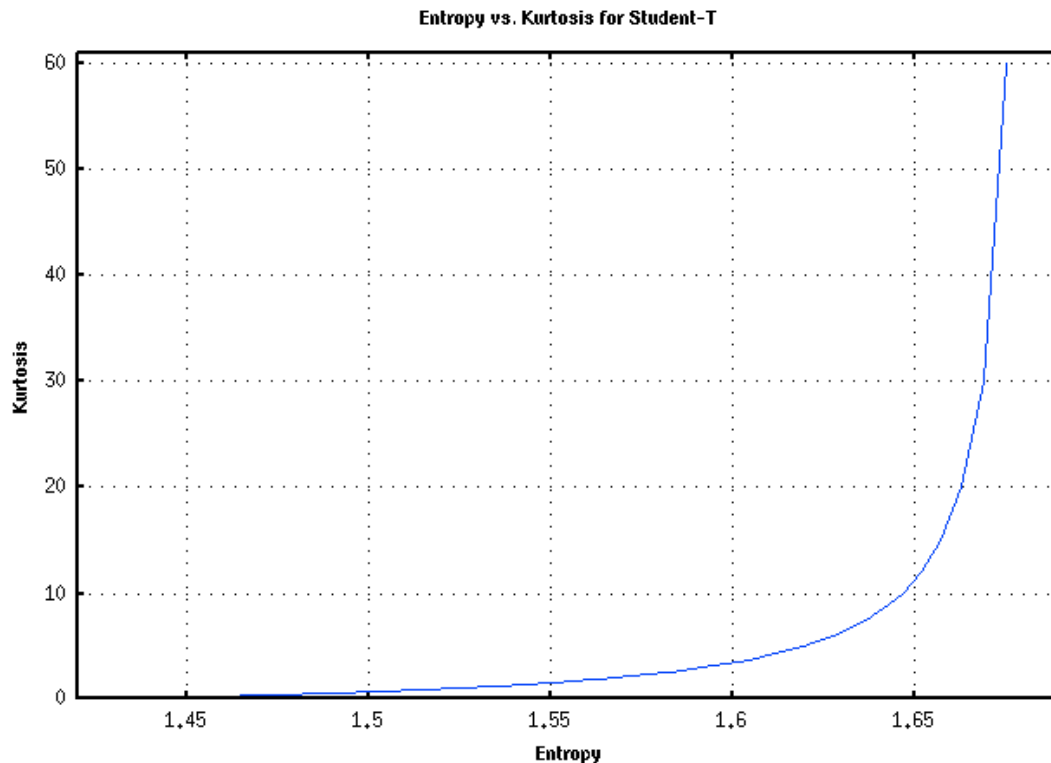


Figure 1: Entropy vs. Kurtosis for the Student's-T distribution.

In the figure above the tail exponent ' ν ' decreases as entropy increases. As ' ν ' tends to infinity, the distribution approaches a gaussian distribution and as ' ν ' drops closer to four it approaches a very fat-tailed distribution with very high excess kurtosis. It is clear the the kurtosis increases as entropy increases, and that a fat-tailed distribution has a higher entropy than a thin-tailed one.

1.1 Rejecting the Gaussian

If we constrain the distribution to be gaussian with a fixed excess kurtosis of zero, then an increase in entropy will see an increase in the variance (second moment). There is no reason for such a constraint to exist however (outside of thermodynamical situations), indeed many distributions in the markets have a kurtosis measure that is well above zero. This rejects the gaussian constraint.

2. Entropy and Information Content

Claude Shannon introduced the notion of entropy as a measure of information content - it quantifies the information contained in a message.

The Efficient Market Hypothesis "asserts that financial markets are "informationally efficient", or that prices on traded assets, e.g., stocks, bonds, or property, already reflect all known information ... " [1] So a time series plot of prices on a traded asset can be seen as containing information about the state of the world. This information is captured by the collective buy and sell actions of traders and each trader adds some information (unique or otherwise) every time she trades.

Markets can then be considered as crucibles of information, just as a hard-disk on a computer stores information - an analog version of a binary stream, if you will.

What happens when more information is allowed into this crucible of information? This increase in information can be due to an increase in transparency, an increase in information discovery and dissipation, an increase in knowledgeable participants, or any other reason - all of which should allow traders to act on new information. The increase in information will be reflected in the time series as an increase in entropy. From section one this implies that the kurtosis of the time series will increase as well.



Figure 2. Information drives markets which absorb the information and respond by varying the entropy and kurtosis.

This allows us to answer the following question: How would the markets respond if company earnings results were released every month instead of every quarter? This hypothesis would suggest an increase in kurtosis on any times-scale over a month. Another concrete example of increased information streaming into the markets is simply an increase in the number of participants in the market. A good proxy for this is liquidity. So the hypothesis would suggest that an increase in liquidity would increase the kurtosis of the market and introduce fatter tails.

The international currency markets are some of the most liquid markets in the world, and operate almost continuously. They also happen to have massively leptokurtic price distributions. The US stock market is one of the most liquid and transparent in the world, while emerging markets are often highly illiquid with limited participants and with limited transparency and information disclosure. US stock markets have very high kurtosis levels compared to emerging markets (emerging markets do have higher second moments).

2.1 Information is the unexpected

If market prices reflect all expected information, then an event that conforms to expectations will add no new information into the markets. On the other hand, an event that is unexpected will add new information to the market (Black Swan). Deviation from expectation is a measure of the change in information. An increase in information content has to be due to an unexpected event. The more unexpected, the higher the increase in information.

So an unexpected event injects a large amount of new information into the system. This increase in information results in an increase in entropy and thus an increase in kurtosis. This is reflected in large fat-tailed moves or an increase in periods of “calm-before-the-storm” behavior.

So the more unexpected an event, the more information it will add, and the larger impact (kurtosis) it will have. And since our expectations are driven by sampling from history, the rarer the event, the more unexpected it will be. Hence the rarer the event, the higher impact it will have (higher kurtosis).

2.2 Non-decreasing Entropy

Consider the time series in Figure 3:

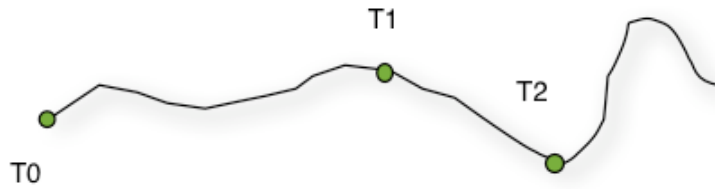


Figure 3. A time-series and its information content

The information content in the period T0-T2 has to be no less than the information content in the period T0-T1 since one encompasses the other. Equality holds if no new information is added between the period T1-T2. Thus:

$$\text{SampleEntropy}(T0-T2) \geq \text{SampleEntropy}(T0-T1)$$

This is the behavior of sample entropy. From Figure 1, distributional kurtosis grows with sample entropy (and vice versa), a similar relationship must hold:

$$\text{DistributionalKurtosis}(T0-T2) \geq \text{DistributionalKurtosis}(T0-T1)$$

Where “ \geq ” is used to denote “Not smaller than” in the two equations above.

This implies distributional kurtosis is non-decreasing as the number of samples increases. But what does the behavior of sample kurtosis look like when distributional kurtosis is non-decreasing?

2.2 Behavior of Sample Kurtosis

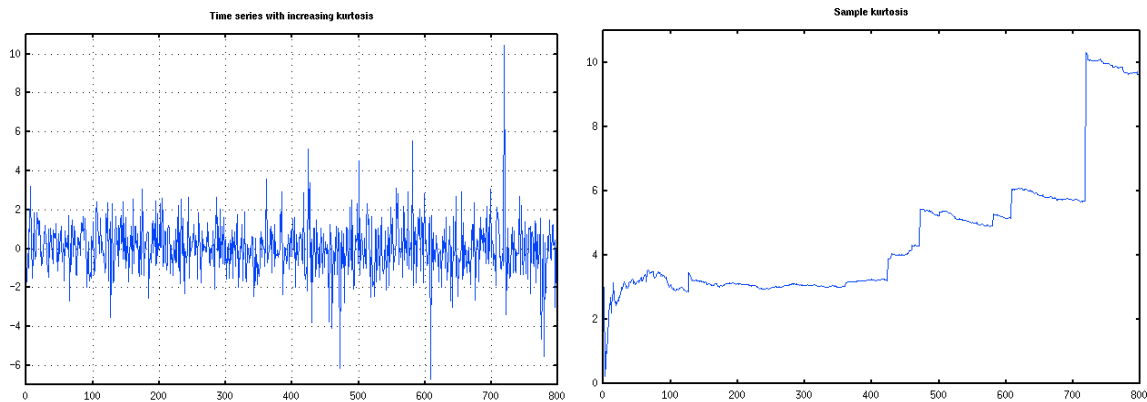


Figure 4. (a) time series with increasing kurtosis and (b) sample kurtosis derived from that time series

Figure 4a is a time series generated with a Student's-T distribution with decreasing tail exponent, and hence increasing kurtosis. It can be seen as a market that is getting more “efficient” with more information streaming in as more informed participants, transparency and liquidity enter the market.

Figure 4b is the sample kurtosis derived from this series with larger and larger intervals starting from T0 (larger and larger sample size). It is clear that the sample kurtosis increases in jumps, with periods where the sample kurtosis actually decreases or remains relatively unchanged interspersed with sudden large contributions from the sample.

So increasing (non-decreasing) *distributional* kurtosis can show decreasing or flat *sample* kurtosis over short periods. However, it will show a net increase in kurtosis in larger time periods. How large is unknown.

High kurtosis distributions can take a very large amount of time to show their properties. The effects are exacerbated by the possibility of “mixed” distributions. Mandelbrot’s original 1963/67 papers [4] has a section on the behavior of sample kurtosis for cotton prices where he shows that the sample kurtosis increases with sample size and tends to infinity as sample size tends to infinity.

Estimation of parameters in the presence of very fat tails is very fragile to estimation errors and the size of the data set needs to be impossibly large to calibrate the tails [5]. We may just not have had enough months (or other large time period) to see the full properties.

3. Testable Predictions

Some testable predictions for the hypothesis:

1. Emerging markets that see an increase in transparency and liquidity (both are easily measured) will see an increase in fat-tailed behavior. The BRIC countries are good candidates to watch.
2. We will see an increase in “calm-before-the-storm” effects. Especially if the notion of volatility continues to be dependent on gaussian tools and used as a risk metric.
3. In countries with economies hinged on innovation and information products, the Gini coefficient of inequality will increase, showing even fatter tails in wealth distribution than at present.
4. Products that increase the discovery, dissipation, and consumption of information will increase fat-tailed behavior (the internet is one example of this, Bloomberg screens, another).
5. As the world tinkers and increases its net knowledge, our actions, driven by this information, will show increasing fat-tailed properties and move further towards extremistan.

4. Notes and References

Note 1: It is also questionable how much of an increase in information 30 extra days can bring in compared to single day. It may be more profitable to compare a single day versus a quarter, since our society has structured information injections on a quarterly basis.

Note 2: There are some indications that larger windows show fatter tails as well [6], but the majority of studies show a decrease in kurtosis for larger time periods.

[1] Efficient Market Hypothesis: http://en.wikipedia.org/wiki/Efficient_market_hypothesis

[2] Jean-Philippe Bouchaud and Marc Potters: <http://www.cfm.fr/us/publications.php>

[3] Nassim Nicholas Taleb, The Fundamental Problem of the 0th Moment: <http://www.fooledbyrandomness.com/zeromoment.pdf>

[4] Mandelbrot, Benoit B., The Variation of Certain Speculative Prices. 1963, 1967. Section 1.8.

[5] Weron, Levy-stable Distributions Revisited: <http://citeseer.ist.psu.edu/448515.html>

[6] Nassim Nicholas Taleb, Poisson Buster and Evidence of Scalability: <http://www.fooledbyrandomness.com/poissonbuster.pdf>