Lesson Notes

MODULE 2 | LESSON 3

JUMPING FOR BETTER VOLATILITY ESTIMATION

Reading Time	30 minutes
Prior Knowledge	Probability density function (pdf), Cumulative density function (cdf), Independent and identically distributed (iid), Value at Risk (VaR), Ito process, Black-Scholes model, Stochastic term, Deterministic term
Keywords	Order statistics, Levy process

In the previous lesson, we learned about variance from the perspective of variance swaps. Specifically, we made the assumption—one of the foundational assumptions of the Black-Scholes model—that security prices follow a Wiener process, meaning they have a drift and a smooth, continuous volatility component. In this lesson, however, we discard that simplifying assumption in favor of an understanding of security price movements that includes non-smooth and non-continuous jumps.

1. Why Jump from the Ito to the Levy Process?

In the first two lessons of this model about variance and volatility, we have made the somewhat simplifying assumption that volatility is itself continuous. This assumption comports well with the assumptions underlying the Black-Scholes equation that is so fundamental to much of what we do in financial engineering, especially as it relates to options theory. We have seen options theory itself applied widely beyond the pricing of vanilla put and call options on equities; options theory has been used in more than one way to model real estate foreclosures, as we saw in the Financial Markets course. By the same token, it can be used to model the probability of default for corporate debt, as in the Merton model. Not to mention, of course, that options theory informs the pricing of more exotic options, including but by no means limited to the variance options we studied in the last two lessons. With all that said, the Black-Scholes model's assumption that price returns behave like an Ito process, with a drift term and a Brownian motion term, may not be the most realistic model for asset prices; it's just a very tractable model that works efficiently and well enough.

The Levy process, on the other hand, can be though of as a generalization of the Ito process from the Black-Scholes model. It also includes the drift term (which can be made deterministic for simplification) and the stochastic component of the continuous Brownian motion. The innovation of the Levy process is the superimposed jump process, generally assumed to itself be a Poisson process. This jump process innovation provides the discontinuity that we refer to as the jump. In the figure below, you see the differential equation that describes the dynamics of the log returns in the model assumed by Spadafora et al. (4). Notice :

1) The continuous volatility component is represented by the dW term in the middle and the dY^c term on the right, where the superscript c stands for "continuous."

2) The dynamics of the jump component are simply represented by the differential term dJ.

3) We mentioned that like the Ito process, the Levy process has a drift term, which may be deterministic. In this case, it is deterministic indeed as it is simply set to zero as assumed by Spadafora et al. and therefore not represented in the equation at all.

$$dY_t = \sigma_t dW_t + dJ_t = dY_t^c + dJ_t$$

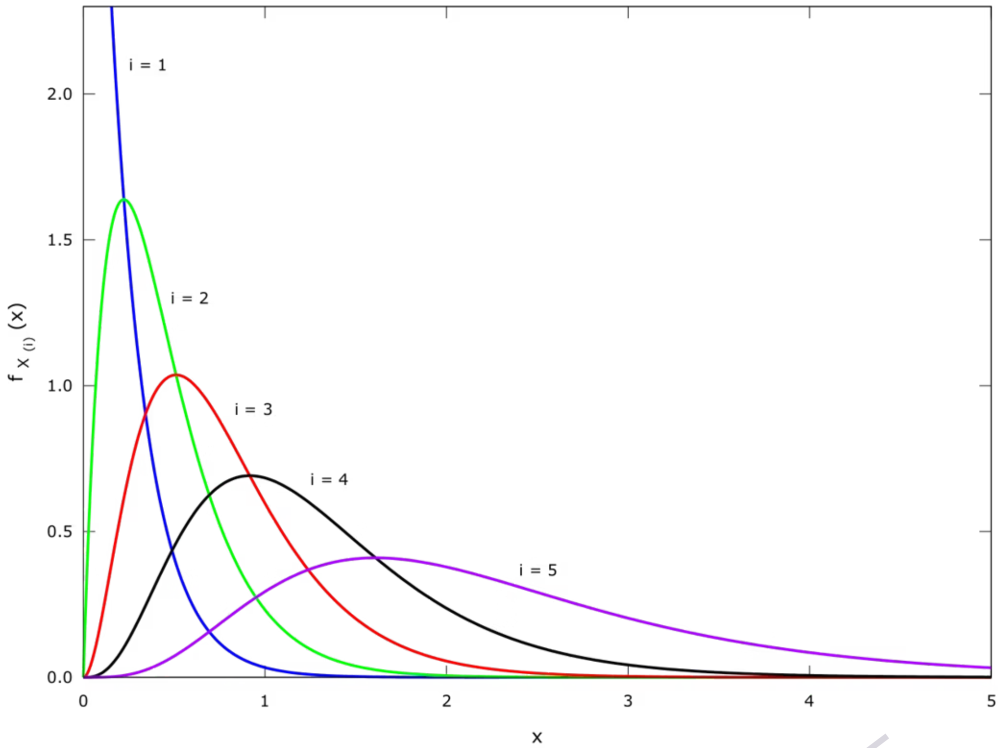
2. Order Statistics and Probability Density Function

In general in order statistics, we want to use what we know about some distribution, especially its cumulative density function (cdf) to make some inferences about a random variable we believe might follow said distribution. So far, that sounds like standard probability and statistics. So more specifically, with order statistics, we are not just interested in the probability that the value of a random number is greater than or less than some given value; rather, we can use order statistics to find the probability that the largest, second largest, third largest, etc., (or smallest, second smallest, or the third smallest) is less than (or greater than) some given value. That is, for a set of, say, *n* independent and identically distributed (iid) observations of a random variable, we can actually construct a probability density function (pdf) for, say, the minimum (lowest value) or maximum (highest value) in this set—by first finding the cdf for the value(s) in question.

As you can see in figure 1, the pdf for the first-order statistic (the smallest value) is furthest to the left (as we would expect) and approaching zero given that zero is the lower end of the support for the exponential distribution. The pdf for the largest value, on the other hand, is furthest to the right: It happens to peak between values 1 and 2 and still has visible probability density at value 5, but the right tail extends asymptotically to infinity. The pdf for the different discrete values indeed overlap, but their peaks are located where we would expect based on what we know about this particular distribution: In this exponential case, with scale parameter equal to one, the pdf is equal to e^(-x).

If we find that, say, for some predefined probability level, the second largest value in a set does not appear to come from the reference distribution we had assumed, we can define that particular observation as a jump.

Figure 1: Probability Density Functions of Order Statistics for sample size n = 5 (from an exponential distribution with unit scale parameter)



Source: "Order Statistic." *Wikipedia*. https://en.wikipedia.org/wiki/Order_statistic.

Given this set of *n* iid observations, we call the smallest the first order statistic, we call the largest the nth order statistic, and we call the k-smallest the kth order statistic (Rundel 2).

Since we have now expanded our scope in this way to include not just the largest (or largest negative, or largest absolute) value, we can make many more inquiries of the empirical data. In the model elaborated by Spadafora et al., the question is not simply, what is the probability that the largest value in our set is a realization from the reference distribution? We now investigate the probabilities that each of the *n* observations is the proper kth order statistic from the reference distribution, with some predefined probability. As Spadafora et al. indicate, “in this way, it could happen that the maximum is not large enough to be classified as a jump,” but the kth maximum can be classified as a jump— despite the fact that it is smaller.

The key insight here from order statistic theory is that the cdf of the kth order statistic of n Standard Normal variables is described as an incomplete beta function as you see in equation 33 of the appendix of the reading, which makes very straightforward the calculation of the probability that an observation belongs to the underlying reference distribution (assumed here to be Gaussian). We will discuss this incomplete beta function in a bit more detail in the next lesson, when we look more closely at the code for the volatility estimator. For now, let's dig into the algorithm that Spadafora et al. use to estimate their volatility estimator based on what we know so far.

3. Integrated Variance and The OS Estimator

1. Let’s now walk through Algorithm 1 to compute the OS estimator.
2. Inputs: As described, we start with the ordered observations of log price returns and compute the naïve sample standard deviation of this sample of observations. We also need to specify our tolerance p, which is the probability we find acceptable that a Gaussian realization is misidentified as a jump.
3. Outputs: Let’s keep in mind what we expect the model to return to us: We want to know the volatility used in the stochastic term of the Ito Process described above. This requires that we effectively filter out the jump observations. These jump observations may be larger than the Gaussian observations, but the authors make it clear that this is not always the case.

4. Step 1a: Compute the threshold θ as described in the paper for the probability p and the two parameters n' and k, where n' is the number of observations assumed to be Gaussian (which initially equals n, the full sample size since we have not identified any observations as jumps yet) and k is the order statistic we are investigating in this iteration. Multiply this, the standard deviation, by this θ, and call this $\hat{\theta}$: if the largest log return is larger than this scaled standard deviation, we consider it a jump, and we subtract one from the number of observations assumed to be Gaussian (in other words, $n' = n' - 1$).
5. Step 2a: Repeat the above except with the absolute value of the largest negative observation. (As we are assuming the reference distribution is Gaussian, we can also assume symmetry.) This time, the standard deviation will only reflect the observation investigated in Step 1a if it was deemed Gaussian. So if need be, recompute $\hat{\theta}$ and compare the absolute value of this largest negative observation to this level. If the absolute value of the observation is larger than this (possibly re-)scaled standard deviation, then we consider this observation a jump and remove it from the set of Gaussian observations we use to calculate the Gaussian volatility.
6. We repeat these steps with the second largest, then the second largest negative observations, and so on, until we have worked our way through the whole sample.

In the latter part of the paper, another algorithm is presented, which extends this model by allowing for heteroskedasticity, or the evolution of the volatility term with time, but relies on the same logic for distinguishing jumps from the continuous (Gaussian) reference distribution.

4. Conclusion

In this lesson, we learned about an extension of the Ito process as a model for the volatility of asset log returns. In this case, in addition to a deterministic and a continuous stochastic term, there is also a jump term. We learned about a model that can take a series of log returns as an input to find the volatility of the continuous stochastic term.

In the next lesson, we see this model in action: We review the the authors' Python implementation of their volatility estimator and experiment with its interpretation of simulated inputs that we design.

References

- "Order Statistic." *Wikipedia*. https://en.wikipedia.org/wiki/Order_statistic.
- Rundel, Colin. "Lecture 15: Order Statistics." Statistics 104, 14 March 2012, Duke University. Lecture. https://www2.stat.duke.edu/courses/Spring12/sta104.1/Lectures/Lec15.pdf
- Spadafora, Luca et al. "Jumping VaR: Order Statistics Volatility Estimator for Jumps Classification and Market Risk Modeling." *arXiv*, 22 Mar 2018, https://arxiv.org/abs/1803.07021.