The Sortino ratio and distributions

December 30, 2019

1 Introduction

The **Sortino ratio** is defined by

$$S = \frac{R - T}{DR},$$

where R is the rate of return of an asset, T is a threshold rate of return, and DR is the **downside variation** of the asset, defined by

$$DR = \sqrt{\int_{-\infty}^{T} (T - r)^2 \cdot f(r)\ dr}.$$

Here, f is the probability density function of the returns. The implicit assumption is that f is integrable, which is a problem for our purposes.

Instead, we only have the empirical cumulative distribution function. For the sequence of observed returns $R_1, R_2, R_3, \ldots$, the nth update of the empirical CDF is given by

$$\widehat{F}_n(x) = n^{-1} \sum_{k=1}^{n} H(x - R_k),$$

where

$$H(x) = \begin{cases} 0 & x < 0 \\ 1 & x \geq 0 \end{cases}$$

is the standard Heaviside function. Unfortunately, H is not differentiable, or even continuous, so $\widehat{f}_n = \widehat{F}_n'$ does not exist.

2 Distributions

Let $\mathcal{D} \subset \mathbb{R}^{\mathbb{R}}$ be a suitably-chosen closed subspace of real functions. We will pin down exactly what constitutes the space once we understand the parameters of our problem. A **distribution** is an element of the dual space $\mathcal{D}^*$ of linear functionals which are continuous with respect to a certain topology. Again, this topology will be determined later. For $G \in \mathcal{D}^*$ and $g \in \mathcal{D}$, we write $\langle G, g \rangle$ to denote the application $G(g)$. Some common distributions include

- evaluation functionals: for $x_0 \in \mathbb{R}$ let $\delta_{x_0} \in \mathcal{D}^*$ be defined by $\langle \delta_{x_0}, g \rangle = g(x_0)$.
- integration functionals: for a suitably chosen function φ, let $I_\varphi \in \mathcal{D}^*$ be defined by $\langle I_\varphi, g \rangle = \int \varphi g$.

We can use integration by parts to motivate a special kind of **distributional derivative**, but this will depend on the specification of $\mathcal{D}$.

Let's examine the downside variation again. To expose the radicand, consider the term

$$DR^2 = \int_{-\infty}^{T} (T-r)^2 \cdot f(r)\ dr.$$

The existence and integrability of the probability density function f is the problem we are trying to fix. We want to turn f into a suitable distribution, which means we will need to identify the correct space of functions $\mathcal{D}$. The only function we know we *need* to be in $\mathcal{D}$ is $r \mapsto (T-r)^2$, which is a $C^\infty((-\infty, T])$ function. Let's run with $\mathcal{D} = C^\infty((-\infty, T])$, and we can topologize it with the seminorms $\{\rho_n\}_{n=1}^{\infty}$ defined by $\rho_n(g) = \|g^{(n)}\|_u$. Thus $g_k \to g$ in $\mathcal{D}$ if and only if $g_k^{(n)} \to g^{(n)}$ uniformly for each n.

With $\mathcal{D}$ specified, we now need to worry about $\mathcal{D}^*$, the space of continuous linear functionals on $\mathcal{D}$. In order for a linear functional ϕ on $\mathcal{D}$ to be continuous, we will need the sequence $\langle \phi, g_k \rangle$ to converge to $\langle \phi, g \rangle$ for every converging sequence $g_k \to g$ in $\mathcal{D}$. For the theory of distributions to be of use to us, then, we need to verify that the Heaviside functions can be viewed as elements in $\mathcal{D}^*$ and confirm that their distributional derivatives are also elements in $\mathcal{D}^*$.

The Heaviside "distribution" H is the functional that operates on $\mathcal{D}$ like $\langle H, g \rangle = \int_0^T g(r)\ dr$. This is because, in the *function* sense, integration against a Heaviside function yields the following equality:

$$\int_{-\infty}^{T} g(r) \cdot H(r)\ dr = \int_{0}^{T} g(r)\ dr.$$

Introduce the shift notation $\langle H_{x_0}, g\rangle = \int_{x_0}^{T} g(r)\ dr$. This again is designed to match the integration intuition:

$$\int_{-\infty}^{T} g(r) \cdot H(r - x_0)\ dr = \int_{x_0}^{T} g(r)\ dr.$$

The set of Heaviside functions $\{H_{x_0}\}_{x_0 \in \mathbb{R}}$ is important for our purposes because the empirical cumulative distribution function $\widehat{F}_n$ is a linear combination of Heavisides. We therefore need to verify that all Heaviside functions reside as distributions in $\mathcal{D}^*$. By appealing to a continuity argument, it will suffice to check that the canonical Heaviside $H = H_0$ is a distribution in $\mathcal{D}^*$. To this end, it is clear that H is linear. To check that H is continuous, suppose $g_k \to g$ in $\mathcal{D}$. On the interval $[0, T]$, each g_k is bounded; moreover, since g_k converges to g uniformly, all but finitely many g_k are bounded above by the function $g + 1$. The dominated convergence theorem therefore implies that $\int_0^T g_k(x)\ dx \to \int_0^T g(x)\ dx$, so $\langle H, g_k\rangle \to \langle H, g\rangle$. This means that H is continuous as a distribution, and so $H \in \mathcal{D}^*$.

Since $\mathcal{D}^*$ is a vector space, and since $H_{x_0} \in \mathcal{D}^*$ for each $x_0 \in \mathbb{R}$, it follows that every empirical cumulative distribution function $\widehat{F}_n$ is in $\mathcal{D}^*$, being a linear combination of Heavisides. The remaining question is whether the distributional derivative of $\widehat{F}_n$ resides in $\mathcal{D}^*$. To this end, it again suffices to verify that $H' \in D^*$, since the rest will follow by a continuity argument. By integration by parts, we have the identity

$$\begin{aligned}
\int_{-\infty}^{T} g(r) \cdot H'(r)\ dr &= g(r) \cdot H(r)\Big|_{-\infty}^{T} - \int_{-\infty}^{T} g'(r) \cdot H(r)\ dr \\
&= g(T) \cdot H(T) - \int_{0}^{T} g'(r)\ dr \\
&= g(T) \cdot H(T) - (g(T) - g(0)) \\
&= g(T) \cdot (H(T) - 1) + g(0),
\end{aligned}$$

so we should define H'_{x_0} to be the distribution $\langle H'_{x_0}, g\rangle = g(T) \cdot (H(T - x_0) - 1) + g(x_0)$. It is not hard to show that H'_{x_0}, so defined, resides in $\mathcal{D}^*$.

3 Conclusion

We can now evaluate $\widehat{F}'_n$. In particular, we have

$$\begin{aligned}
\langle \widehat{F}'_n, g\rangle &= n^{-1}\sum_{k=1}^{n}\langle H'_{R_k}, g\rangle \\
&= n^{-1}\sum_{k=1}^{n}\left[g(T)\cdot(H(T-R_k)-1)+g(R_k)\right] \\
&= n^{-1}\cdot g(T)\sum_{k=1}^{n}\left[H(T-R_k)-1\right]+n^{-1}\sum_{k=1}^{n}g(R_k).
\end{aligned}$$

This is the evaluation on an arbitrary $g \in \mathcal{D}$. For the specific function $g(r) = (T-r)^2$ in the Sortino ratio, this becomes

$$\langle \widehat{F}'_n, g\rangle = n^{-1}\sum_{k=1}^{n}(T-R_k)^2$$

since $g(T) = 0$. The nth estimate of the Sortino ratio is, therefore, given by

$$\widehat{S}_n = \frac{R_n - T}{\widehat{DR}_n} = \frac{R_n - T}{\sqrt{n^{-1}\sum_{k=1}^{n}(T-R_k)^2}}.$$

Interestingly, and perhaps coincidentally, the computation of $\langle \widehat{F}'_n, g\rangle$ corresponds to the uniform Monte Carlo sampling estimate of the original integral.