The Sortino ratio and distributions

January 1, 2020

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 $r_0 \in \mathbb{R}$ let $\delta_{r_0} \in \mathcal{D}^*$ be defined by $\langle \delta_{r_0}, g\rangle = g(r_0)$.
- 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. On the other hand, $r \mapsto (T-r)^2$ doesn't have the right asymptotic properties on which to assign a topology of uniform convergence directly. This can be resolved by introducing a weight function $w(r) = (T-r)^{-2} + 1$ and by defining $\mathcal{D}$ to be the subspace of $C^\infty((-\infty, T])$ given by

$$\mathcal{D} = \{g \in C^\infty((-\infty, T]) : \ \|g^{(n)}\|_w < \infty \text{ for all } n \in \mathbb{N}\},$$

where

$$\|g\|_w = \sup_{r\in\mathbb{R}}\{((T-r)^{-2}+1)|g(r)|\}.$$

Clearly, $r \mapsto (T-r)^2$ is a member of $\mathcal{D}$, and there is a natural choice of topology with the seminorms $\{\rho_n\}_{n=1}^\infty$ defined by $\rho_n(g) = \|g^{(n)}\|_w$. Thus $g_k \to g$ in $\mathcal{D}$ if and only if $w \cdot g_k^{(n)} \to w \cdot g^{(n)}$ uniformly for each n. This, in turn, implies that $g_k^{(n)} \to g^{(n)}$ uniformly on compact sets 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 D on $\mathcal{D}$ to be continuous, we will need the sequence $\langle D, g_k\rangle$ to converge to $\langle D, g\rangle$ 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 $\mathcal{D}^*$.

Define the Heaviside functional H_{r_0} by

$$\langle H_{r_0}, g\rangle = H(T-r_0)\int_{r_0}^{T} g(r)\, dr.$$

The term $H(T - r_0)$ is 1 if $r_0 \le T$ and 0 otherwise. This ensures that the Heaviside functional is appropriately 0 when the center r_0 is to the right of T. We first verify that H_{r_0} is linear, then verify its continuity. Indeed,

$$\begin{aligned}\langle H_{r_0}, c_1 g_1 + c_2 g_2 \rangle &= H(T - r_0) \cdot \int_0^T c_1 g_1(r) + c_2 g_2(r) \; dr \\ &= c_1 H(T - r_0) \int_{r_0}^T g_1(r) \; dr + c_2 H(T - r_0) \int_{r_0}^T g_2(r) \; dr \\ &= c_1 \langle H_{r_0}, g_1 \rangle + c_2 \langle H_{r_0}, g_2 \rangle,\end{aligned}$$

establishing linearity. Moreover, if $g_k \to g$ in $\mathcal{D}$, then $g_k \to g$ uniformly on $[r_0, T]$. Therefore, all but finitely many g_k are bounded above by $g+1$; hence, $\int_{r_0}^T g_k \to \int_{r_0}^T g$ by the dominated convergence theorem. This implies that H_{r_0} is a continuous function. Therefore, $H \in \mathcal{D}^*$.

Since $\mathcal{D}^*$ is a vector space, and since $H_{r_0} \in \mathcal{D}^*$ for each $r_0 \in \mathbb{R}$, it follows that every empirical CDF

$$\widehat{F}_n = n^{-1} \sum_{k=1}^n H_{R_k}$$

is in $\mathcal{D}^*$, being a linear combination of Heavisides. The remaining question is whether the distributional derivative of $\widehat{F}_n$ resides in $\mathcal{D}^*$. It suffices to consider whether $H'_{R_0} \in \mathcal{D}^*$. We needs first agree to a suitable *notion* of differentiation. Imagining a successful application of integration by parts with suitable functions, we would expect that

$$\begin{aligned}\int_{-\infty}^T g(r) \cdot H'(r - r_0) \; dr &= g(r) \cdot H(r - r_0) \Big|_{-\infty}^T - \int_{-\infty}^T g'(r) \cdot H(r - r_0) \; dr \\ &= g(T) \cdot H(T - r_0) - \int_{r_0}^T g'(r) \; dr \\ &= g(T) \cdot H(T - r_0) - (g(T) - g(r_0)) \\ &= g(T) \cdot [H(T - r_0) - 1] + g(r_0).\end{aligned}$$

For consistency, we *define* H'_{r_0} by

$$\langle H'_{r_0}, g \rangle = H(T - r_0) \cdot [H(T - r_0) - 1] \cdot g(T) + H(T - r_0) \cdot g(r_0).$$

It is clear that H'_{r_0} is linear and continuous, so $H'_{r_0} \in \mathcal{D}^*$. Therefore $\widehat{F}'_n \in \mathcal{D}^*$ for each $n \in \mathbb{N}$.

3 Conclusion

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

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

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

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

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

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

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.