Econometric Reviews ›
Volume 27, 2008 - Issue 4-6

Original Articles

Generalized Safety First and a New Twist on Portfolio Performance

M. Ryan Haley & Charles H. Whiteman

Pages 457-483 | Received 14 Nov 2005, Accepted 08 Nov 2006, Published online: 22 May 2008

Cite this article https://doi.org/10.1080/07474930801960360

Abstract

We propose a Generalization of Roy's (1952) Safety First (SF) principle and relate it to the IID versions of Stutzer's (Stutzer's 2000, 2003) Portfolio Performance Index and underperformance probability Decay-Rate Maximization criteria. Like the original SF, the Generalized Safety First (GSF) rule seeks to minimize an upper bound on the probability of ruin (or shortfall, more generally) in a single drawing from a return distribution. We show that this upper bound coincides with what Stutzer showed will maximiz ge return shrinks t is simple enough Deviations theo ond to the Kullba and the "closest" n which minimizi lead to the same de nus that the

decay-rate maximizing strategy may require the investor to take positions that do not minimize the probability of shortfall in each successive period. It also makes clear that the relationship between the marginal distribution of the one-period portfolio return and the mean-shortfall distribution is the same as that between the source density and the target density in importance sampling. Thus Geweke's (1989) measure of Relative Numerical Efficiency can be used as a measure of the quality of the divergence measure. Our interpretation of the decay rate maximizing criterion in terms of a one-shot problem enables us to use the tools of importance sampling to develop a "performance index" (standard error) for the Portfolio Performance Index (PPI). It turns out that in a simple stock portfolio example, portfolios within one (divergence) standard error of one another can have very different weights on individual securities.

Keywords: Entropy | Importance sampling | Kullback–Leibler divergence | Portfolio choice | Portfolio performance | Safety first | Shortfall

JEL Classification: G11 | C4

ACKNOWLEDGMENTS

The article was prepared for the 2nd Conference on Information and Entropy in Econometrics, American University, September 23–25, 2005. We are grateful to Joel L. Horowitz, Narayana Kocherlakota, Michael J. Stutzer, and two anonymous referees for helpful comments and useful suggestions.

Notes

[1] R(w) ca… cause our analysis … return.

[2] In relat… places the indic… ints out that similar p…

[3] Stutzer… dependent data. Th… case that

the returns are IID, though a similar rate function characterizes the non-IID case. Since much of our analysis hinges on the precise form of I(d,w), it should be regarded as applying only to the IID case. We conjecture that something very similar to our analysis would apply in the non-IID case.

[4]In the log-optimal version in Stutzer (2003), $1 - \theta$ is interpreted as the coefficient of relative risk aversion.

[5]An additional interpretation of θ as a Lagrange multiplier will be offered in the next subsection.

[6]Technically, there is an important distinction between a divergence and a distance; the former is not a proper metric and may violate properties such as symmetry or the triangularity rule.

[7]For a general proof of Kullback's lemma see, for example, Bucklew (1990, p. 30).

[8]Note that the probabilities, $\pi_t(\cdot)$, are concentrated in terms of the to-be-determined multiplier θ; this reduces the dimensionality of the optimization problem from $(T + N + 1)$ to $(N + 1)$. For a more thorough discussion about the relationship between the $\pi_t(\cdot)$s, I(d,w), GSF, and the KL divergence, see Haley (2003).

[9]To match the conditions in Stutzer (2003), the mean of X should be strictly greater than d to ensure that the probability of shortfall goes to zero asymptotically. Our argument is clearer with $E(X) = d$; continuity ensures that it will go through with a slightly larger mean.

[10]Strictly speaking, the “infinitely-repeated” terminology applies only in the case that returns are IID. In the general non-IID case treated in Stutzer (2003), the term “dynamic game” is more appropriate.

[11]Horizo ze discount ometric.

[12]In re er using the sum- e.

[13]That is e PPI. Such measure imated

moments or parameters, as sampling error will affect the performance of each such procedure.

[14]This notation generally follows that of Geweke (1989). For simplicity, we take $\mathcal{I}(\psi)$ and $p(\psi)$ to be proper normalized densities; Geweke works with the more general unnormalized (kernel) density.

[15]The adjective “numerical” is used to emphasize that even in a fully Bayesian context, frequentist procedures may be appropriate for assessing the sampling properties of a posterior sample generated randomly using Monte Carlo procedures. We will apply the same reasoning to the data sample, so the standard terminology applies.

∗Sample size equals 240.

[16]Two stocks Stutzer (2000) used have dropped out of the CRSP data set.

∗Sample size equals 240.

∗Sample size equals 240.

[17]The classic example is the “height with shoes on vs. height with shoes off” example: the population variation in heights is irrelevant, as everyone is taller with shoes on.