







Q

Home ► All Journals ► Mathematics, Statistics & Data Science ► Econometric Reviews ► List of Issues ► Volume 27, Issue 4-6 ► Generalized Safety First and a New Twist

Econometric Reviews >

Volume 27, 2008 - Issue 4-6

170 22 0
Views CrossRef citations to date Altmetric

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



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 maximize the rate at which the probability of shortfall in the long-run average return shrinks to zero in repeated drawings from the return distribution. Our setup is simple enough that we can illustrate via direct calculation a deep result from Large Deviations theory: in the IID case the GSF probability bound and the decay rate correspond to the Kullback-Leibler (KL) divergence between the one-shot portfolio distribution and the "closest" mean-shortfall distribution. This enables us to produce examples in which

minimizing the upper bound on the underperformance probability does not lead to the same decision as minimizing the underperformance probability itself, and thus 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:



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) can be interpreted as either a net return or a logged gross return; because our analysis applies to both cases equally, we use refer to R(\bf w) simply as a return.

²In related work (Haley et al., <u>2007</u>), we are exploring an alternative that replaces the indicator function with a smooth, differentiable approximation. A referee points out that similar problems and solutions arise in the literature on machine learning.

³Stutzer (2003) is careful to point out how matters change with temporally dependent data. The I(d,w) function we work with describes the decay rate only in the 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.

⁴In 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.

⁶Technically, 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.

⁷For a general proof of Kullback's lemma see, for example, Bucklew (<u>1990</u>, p. 30).

⁸Note 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 ($\underline{2003}$), 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.

¹⁰Strictly 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.

¹¹Horizon dependence is of course not unique to the PPI, and will characterize discounted expected utility procedures whenever the discounting is not geometric.

¹²In related work, Haley and McGee (2006) explore these relationships further using the sum-of-squared deviations measure of disparity in place of the KL divergence.

¹³That is, we seek a measure of the effect sampling error might have on the PPI. Such measures are relevant for all portfolio allocation procedures that rely on estimated moments or parameters, as sampling error will affect the performance of each such procedure.

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

¹⁵The 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.

¹⁷The 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.

Related Research Data

A Portfolio Performance Index

Source: Financial Analysts Journal

Probability of Survival as an Investment Criterion

Source: Management Science

The Sharpe Ratio

Source: The Journal of Portfolio Management

Risk and Rank or Safety First Generalised

Source: Economica

Myopic Loss Aversion and the Equity Premium Puzzle

Source: The Quarterly Journal of Economics Safety First--An Expected Utility Principle

Source: Journal of Financial and Quantitative Analysis

A Simple Nonparametric Approach to Derivative Security Valuation

Source: The Journal of Finance

Related research 1



People also read Recommended articles Cited by 22

Information for

Authors

R&D professionals

Editors

Librarians

Societies

Opportunities

Reprints and e-prints

Advertising solutions

Accelerated publication

Corporate access solutions

Open access

Overview

Open journals

Open Select

Dove Medical Press

F1000Research

Help and information

Help and contact

Newsroom

All journals

Books

Keep up to date

Register to receive personalised research and resources by email



Sign me up











Accessibility



Copyright © 2025 Informa UK Limited Privacy policy Cookies Terms & conditions



Registered in England & Wales No. 01072954 5 Howick Place | London | SW1P 1WG