A Sharper Ratio: A General Measure for Correctly Ranking Non-Normal Investment Risks

Kent Smetters *　　　Xingtan Zhang †

This Version: February 3, 2014

Abstract

While the Sharpe ratio is still the dominant measure for ranking risky investments, much effort has been made over the past three decades to find more robust measures that accommodate non-Normal risks (e.g., "fat tails"). But these measures have failed to map to the actual investor problem except under strong restrictions; numerous ad-hoc measures have arisen to fill the void. We derive a generalized ranking measure that correctly ranks risks relative to the original investor problem for a broad utility-and-probability space. Like the Sharpe ratio, the generalized measure maintains wealth separation for the broad HARA utility class. The generalized measure can also correctly rank risks following different probability distributions, making it a foundation for multi-asset class optimization. This paper also explores the theoretical foundations of risk ranking, including proving a key impossibility theorem: any ranking measure that is valid for non-Normal distributions cannot generically be free from investor preferences. Finally, we show that approximation measures, which have sometimes been used in the past, fail to closely approximate the generalized ratio, even if those approximations are extended to an infinite number of higher moments.

Keywords: Sharpe Ratio, portfolio ranking, infinitely divisible distributions, generalized ranking measure, Maclaurin expansions
JEL Code: G11

*Kent Smetters: Professor, The Wharton School at The University of Pennsylvania, Faculty Research Associate at the NBER, and affiliated faculty member of the Penn Graduate Group in Applied Mathematics and Computational Science. By Email: `smetters@wharton.upenn.edu`.

†Xingtan Zhang, PhD (Penn Mathematics), and first-year PhD student in Applied Economics at The Wharton School at The University of Pennsylvania. By email: `xingtan@wharton.upenn.edu`.

1 Introduction

Bill Sharpe's seminal 1966 paper demonstrated that picking a portfolio with the largest expected risk premium relative to its standard deviation is equivalent to picking the portfolio that maximizes the original investor's expected utility problem, assuming that portfolio returns are Normally distributed.[1] This simple mean-variance investment ranking measure — the "Sharpe ratio" — is, therefore, a sufficient statistic for the investor's problem that does not rely on the investor's preferences or wealth level.

The immense power of the Sharpe ratio ranking measure stems from the fact that it allows the investment management process to be decoupled from the specific attributes of the heterogeneous investor base. The multi-trillion dollar money management industry relies heavily on this separation. Investors in a mutual fund or hedge fund might differ in their levels of risk aversion and wealth (including assets held outside the fund). Nonetheless, an investment manager only needs to correctly estimate the first two moments of the fund's return in order to pick the single risky portfolio that is best for each underlying investor.[2] It is not surprising, therefore, that the Sharpe ratio is tightly integrated into the modern investment management practice and embedded into virtually all institutional investment analytic and trading platforms. Even consumer-facing investment websites like Google Finance reports the Sharpe ratio for most mutual funds along with just a few other basic statistics, including the fund's alpha, beta, expected return, R^2 tracking (if an indexed fund), and standard deviation.

Of course, it is well known that investment returns often exhibit "higher order" moments that might differ from Normality (Fama 1965; Brooks and Kat 2002; Agarwal and Naik 2004, and Malkiel and Saha 2005).[3] In practice, investment professionals, therefore, often look for investment opportunities that would have historically — that is, in a "back test" — produced unusually large Sharpe ratios under the belief that large values provide some "buffer room" in case the underlying distribution is not Normal. This convention, though, is misguided. Outside of the admissible utility-probability space (*"admissible space"* for short) where the Sharpe ratio is valid, it is easy to create portfolios with large Sharpe ratios that are actually first-order stochastically *dominated* by portfolios with smaller Sharpe ratios (Leland 1999; Spurgin 2001; and Ingersoll et al. 2007).[4] Besides the presence of the usual "fat tails," it is now well known that non-Normally distributed risks easily emerge at the investment fund level with modern trading strategies and the use of financial derivatives (Section 6), which have exploded in use over time.

The historical debate over the ability of the Sharpe ratio to correctly rank risky investments following non-Normal probability distributions (e.g., Tobin 1958, 1969; Borch 1969; Feldstein 1969) has led to a long-standing interest in producing more robust ranking measures. The first line of work dates back to at least the work by Paul Samuelson (1970), who, at the time, also expressed some skepticism that such extensions were actually needed in practice. (The subsequent explosion of modern trading strategies, novel asset classes and financial derivatives might have changed his mind.) A short list of other contributors include the well-cited paper by Kraus and Litzenberger (1976), Scott and Horvath (1980), Owen and Rabinovitch (1983), Ingersoll's (1987) classic textbook, Brandt et al (2005), Jurczendo and

[1]Of course, the Sharpe ratio builds on the pioneering mean-variance work by Markowitz (1952, 1959) and Tobin (1958).

[2]The Sharpe ratio, however, only ranks *risky* portfolios in order to determine the best one. The ratio itself does not determine the optimal division of an investor's wealth between this best risky portfolio and the risk-free instrument. That division must be determined in a second stage using consumer-specific information. The Sharpe ratio, therefore, supports the standard division between the "investment manager," who determines the best risky portfolio, and the "financial planner" who, knowing each client in more detail, helps decide the share of the client's wealth that should be invested into this best risky asset based on the client's specific circumstances.

[3]A related literature has examined how disaster risk can explain equilibrium pricing within the neoclassical growth model (Barro 2009; Gabaix 2012; Gourio 2012; Wachter 2012)

[4]In other words, the portfolio with the *smaller* Sharpe ratio would be preferred by all expected utility maximizers with positive marginal utility in wealth.

Maillet (2006), Zakamouline and Koekebakker (2008), Dávila (2010) and Pierro and Mosevich (2011). This line of work, however, imposed fairly strong restrictions on investor utility preferences and/or the risk distribution. The current paper contributes to this line of research by deriving a ranking measure that is valid over a broad admissible space.[5]

A second line of research bypasses the investor's expected utility problem altogether and produces risk measures that satisfy certain mathematical properties such as "coherence."[6] Examples of coherent risk measures include "average VaR," "entropic VaR," and the "superhedging price." While these measures satisfy certain axioms, a portfolio that maximizes one or more of these measures does not necessarily maximize the standard investor expected utility problem, as considered by Sharpe and many others. The application of these measures for the actual investor is, therefore, unclear, which might help explain the continued popularity of the Sharpe ratio.

A third line of work, which is actually the largest line in scope, has evolved more from practitioners. It has produced heuristic measures that have a more "intuitive" interpretation in nature than the axiomatic-based measures. Common heuristic measures include "value at risk (VaR),"[7] Omega, the Sortino ratio, the Treynor ratio, Jensen's alpha, Calmar ratio, Kappa, Roy's safety-first criterion, numerous tail risk measures, various upside-downside capture metrics, and many more.[8] These metrics, however, tend to be especially problematic. Not only do they fail to satisfy any sort of reasonably mathematical properties, there is no apparent relationship between a reasonable description of the investor problem and these measures. In practice, therefore, investment managers often combine the Sharpe ratio with one or more of these measures when attempting to account for non-Normal risk (e.g., maximize the Sharpe ratio subject to the investment's "value at risk" being less than some threshold). Despite its limitations, the Sharpe ratio, therefore, remains the gold standard of the investment industry.

This paper makes three contributions. *First*, as summarized in our Lemma 2, we demonstrate how to solve an infinite-order Maclaurin expansion for its correct asymptotic root when no closed form solution exists. We can then derive a generalized ranking measure (the "generalized ratio") that correctly ranks risky returns under a much broader admissible utility-probability space consistent with the Sharpe ratio or previous extensions. By "correctly ranks," we mean it in the tradition of Sharpe: the generalized ratio picks the portfolio preferred by the original investor expected utility problem.

It is easy to motivate the importance of allowing for a broad admissible utility-probability space. A broad utility space captures realistic investor attitudes toward risk. For example, while the common assumption of Constant Absolute Risk Aversion is useful for obtaining various theoretical insights, it is also fairly implausible for modeling risky investment decisions. Similarly, allowing for a broad set of risk distributions is, of course, important for accommodating more extreme risks with "fat tails."

But our generalized ratio can also rank between risks that follow *different* probability distributions. The generalized ratio, therefore, can be used as the foundation for multi-asset class optimization. For example, it can pairwise rank a risky portfolio without financial derivatives that follows one distribution

[5]Throughout this paper, we will write expressions like "ranking measure ABC is valid over admissible space XYZ" even though such terminology is a bit redundant since admissibility implies validity. However, we believe that such terminology is generally understood and more readable than various alternatives.

[6]A "coherent" risk measure satisfies monotonicity, sub-additivity, homogeneity, and translational invariance (Artzner et al 1999). More recent work has emphasized risk measures that avoid "worst case" scenarios and are monotonic in first-order stochastic dominance. See, for example, Aumann and Serrano (2008); Foster and Hart (2009); and Hart (2011).

[7]Standard VaR is not coherent, whereas the variants on VaR noted in the previous paragraph are coherent.

[8]Modigliani (1997) proposed a transformation of the Sharpe ratio, which became known as the "risk-adjusted performance measure." This measure attempts to characterize how well a risk rewards the investor for the amount of risk taken relative to a benchmark portfolio and the risk-free rate. This measure is not included in the list in the text because it mainly provides a way of interpreting the unit-free Sharpe ratio rather than offering an alternative measure in the presence of non-Normally distributed risk.

against another risky portfolio with option overlays following a different distribution. This flexibility is much more powerful than simply assuming that all potential portfolio combinations follow the same probability distribution form, even if that distribution is more extensible than Normality.

Like the original Sharpe ratio, our generalized measure preserves wealth separation under the broad functional form of HARA utility, which includes many standard utility functions as special cases.[9] Unlike the original Sharpe ratio, however, our generalized ratio does not preserve separation from investor preferences. But we show that this limitation is not a function of the generalized risk measure. Rather, we prove a key impossibility theorem: preference separation is generically impossible in the presence of non-Normal risk. Fortunately, as a practical matter, the generalized ratio still supports the decoupled investment management process noted above: instead of reporting a single ratio, each fund can report a small tuple of ratios corresponding to different standardized levels of risk aversion (see Section 6). By law, financial advisors must already actively test for the level of risk aversion of each client.

Second, using some of the machinery that we developed, we then "backtrack" to explore the theoretical foundations of the classic Sharpe ratio in more detail. Despite its extensive usage in academics and industry, very little is actually known about the Sharpe ratio beyond the few cases where it is well known to correctly rank risks (e.g., Normally distributed risk or quadratic utility). We show that the Sharpe ratio is actually valid under a larger admissible space than currently understood. We also explore why it is challenging to actually write down a *necessary* condition for the Sharpe ratio to be a valid ranking measure. In the process, we are also able to generalize the Kraus-Litzenberger (1976) "preference for skewness" result to an unlimited number of higher moments. This generalization is useful because plausible utility functions produce an infinite number of non-zero higher-order derivatives, and there does not exist any probability distribution that can be fully described by any *finite* number of cumulants greater than two.

Third, we derive a linear approximation of the investor problem in the presence of non-Normal higher-order moments. This formulation accommodates a simple closed-form solution, and it nests some previous attempts to generalize the Sharpe ratio. Our computations, however, show that approximations can be very inaccurate. Accurate ranking, therefore, requires using the generalized ratio.

The paper is organized as follows. Section 2 provides an overview of the standard investor problem. Section 3 derives the generalized ratio described earlier. Section 4 explores the theoretical foundations of the Sharpe ratio in more detail. Section 5 derives the linear approximation. Section 6 provides numerical examples comparing the Sharpe ratio, the generalized ratio, and the linear approximation for a range of potential investment applications. Section 7 concludes. Proofs of lemmas and theorems are provided in the Appendices.

2 The Investor Problem and the Sharpe Ratio

2.1 Investor Problem

The investor has preferences characterized by the utility function u and wants to allocate wealth w among the risk free asset paying a constant rate r and a risky asset paying a net return Y. More formally:

$$\max_a \mathbb{E}u\left(w(1+r)+a(Y-r)\right) \tag{1}$$

where the variable a is the amount of wealth invested in the risky asset. To reduce notation, we will often write $w_r \equiv w(1+r)$. Now suppose that u belongs to the function space $\mathscr{U}_s$ that denotes all the

[9] In other words, the generalized ratio can correctly rank without knowing the investor's wealth.

smooth utility functions defined on the real number line with positive odd-order derivatives and negative even-order derivatives. $\mathscr{U}_s$, therefore, incorporates a broad set of utility classes including HARA. Of course, this problem may not be well defined if, for example, the utility function is only defined for positive wealth while the risk distribution is unbounded. So, as usual, we restrict the *admissible space*, defined more rigorously below, to well defined utility-probability combinations.

Lemma 1. *For any given increasing and concave utility function u, maximization problem (1) has a unique solution a^*. Furthermore,*
if $\mathbb{E}Y > r$ then $a^* > 0$;
if $\mathbb{E}Y = r$ then $a^* = 0$;
if $\mathbb{E}Y < r$ then $a^* < 0$

In other words, the investor problem that we are considering is standard. A unique best portfolio exists that maximizes the investor's expected utility. Moreover, risk taking follows the usual behavior: if the risky security's expected return exceeds the risk-free rate then some of the risk will be held; if the two returns are equal then no risk is held; otherwise, a short position is taken in the risky asset.

Given the utility function u and initial wealth w, then for two different risky assets, Y_1 and Y_2, and risk free rate r, it is a convenient shorthand to write $(Y_{1,}r) \geq_u^w (Y_{2,}r)$ if and only if

$$\max_a \mathbb{E}u(w(1+r)+a(Y_1-r)) \geq \max_a \mathbb{E}u(w(1+r)+a(Y_2-r))$$

In words, $(Y_{1,}r) \geq_u^w (Y_{2,}r)$ implies that an investor with preference function u and initial wealth w prefers investing in Y_1 over Y_2 when the risk-free is r. (We will sometimes write $(Y_{1,}r) \geq_u (Y_{2,}r)$ if $(Y_{1,}r) \geq_u^w (Y_{2,}r)$ for all $w > 0$.) Note that Y_1 and Y_2 can be risks derived from different probability distributions.

2.2 Ranking Definitions

Like the original Sharpe Ratio, we want to pairwise rank two risky investments with random returns Y_1 and Y_2. Of course, if we know the investor's wealth, preferences and the exact form of the underlying risk distribution, we can then simply integrate the expectation operator in equation (1) to determine investor's indirect utility associated with each risk. However, in practice, we are typically missing some of this information, and so we would like to be able to rank investments based on a subset of this information. Indeed, as noted in Section 1, the real power of the Sharpe Ratio stems from its ability to correctly pairwise rank two investment risks simply by knowing the first two moments of the underlying Normal distribution and the risk-free rate. Toward that end, the following definitions are useful:

Definition 1. [Ranking Measures] For any risky asset Y and risk-free rate r, we say that:

- A *distribution-only ranking measure* is a function q_D which only depends on Y and r.
- A *distribution-preference ranking measure* is a function q_{DU} which only depends on Y, r, and u.
- A *distribution-preference-wealth ranking measure* is a function q_{DUW} which only depends on Y, r, u and w.

Definition 2. [Valid Ranking Measure, Admissible Space] Suppose $\mathscr{U}$ is a set of utility functions and $\mathscr{Y}$ is a set of random variables. We then say that q is a *valid* ranking measure with respect to $\mathscr{U} \times \mathscr{Y}$ if, $\forall u \in \mathscr{U}$ and $\forall Y_{1,}Y_2 \in \mathscr{Y}$:

$$q(Y_1, r, \bullet) \geq q(Y_2, r, \bullet) \Longleftrightarrow (Y_{1,}r) \geq_u^w (Y_{2,}r)$$

We call $\mathscr{A} \equiv \mathscr{U} \times \mathscr{Y}$ the *admissible space* of the ranking measure. We can also define q_n to be a valid ranking measure *sequence* if it point-wise converges to a valid ranking measure.

In words, a valid *distribution-only* ranking measure only requires knowing the properties of the risk distribution, and not the investor's preferences or level of wealth, in order to properly rank risks. The Sharpe ratio is an example. A valid *distribution-preference* ranking measure then also requires knowing the investor's preferences. We show below that our generalized ranking measure for HARA utility is an example. We also prove that all ranking measures that are valid at arbitrary higher moments of the risk distribution and across a wide range of utility functions must at least be a distribution-preference ranking measure. Finally, the *distribution-preference-wealth* ranking measure then requires also knowing the investor's wealth. Of course, this ranking measure is the least powerful of the three. Relative to the original investor problem, its main advantage is that it does not require knowing the exact risk distribution: valid ranking can still be achieved by estimating the distribution's moments empirically.

Importantly, set $\mathscr{Y}$ can include many risk distribution families. That flexibility allows comparing portfolios with potentially very different assets. Of course, not all utility-risk distribution pairs are compatible with a well-defined investor problem described by Lemma 1. For example, a risk distribution with unbounded negative returns can't be combined with a utility function where the Inada condition holds ($\lim_{w\to 0} u'(w) = \infty$). These combinations are implicitly ruled out in Definition 2 through the mapping back to the original investor problem (1).

Remark 1. It is straightforward to show that a sufficient condition for function q to be valid ranking measure is for the indirect utility, $\max_a \mathbb{E}u(w(1+r)+a(Y-r))$, to be increasing in q. Moreover, two ranking measures are equivalent if one measure is a strictly increasing transformation of the other.

Example 1. Let $\mathscr{U}_e = \{u(\cdot) : u(w) = -exp(-\gamma w), \gamma > 0\}$ be a set of utility functions, commonly known as the CARA class. Also, let $q_{CARA}(Y,r) = \gamma\{u[\max_a \mathbb{E}u(w(1+r)+a(Y-r))] - w(1+r)\}$. Then, it is easy to show that $q_{CARA}(Y,r)$ only depends on Y and r for $\forall u \in \mathscr{U}_e$ and is a valid distribution-only ranking measure with respect to $\mathscr{U}_e \times \mathscr{Y}$, where $\mathscr{Y}$ is the set of all random variables.

2.3 The Sharpe Ratio

Suppose that risky return Y is drawn from a Normal distribution $N\left(\mathbb{E}Y, \sqrt{Var(Y)}\right)$, and let

$$q_S(Y,r) = \left(\frac{\mathbb{E}Y - r}{\sqrt{Var(Y)}}\right)^2.$$

Sharpe (1966) showed that the investor's indirect utility (1) is an increasing function of $q_S(Y,r)$, the Sharpe ratio squared. (If, in addition, we restrict the random variable space so that $\mathbb{E}Y \geq r$, which generally holds in equilibrium, then the investor's indirect utility is an increasing function of the more familiar form of the Sharpe ratio, $\frac{\mathbb{E}Y - r}{\sqrt{Var(Y)}}$.) Hence, q_S is a valid distribution-only ranking measure with respect to set $\mathscr{Y}_n$ of all Normal distributions. It can also be shown that q_S is a valid distribution-only ranking measure with respect to the set $\mathscr{U}_q$ of all quadratic utility functions.

However, the Normal distribution is only a sufficient condition for q_S to be a valid distribution-only based ranking measure. The Sharpe ratio ranking measure q_S is actually valid over a wider class of return distributions.

Theorem 1. *Let χ_α denote a parametrized distribution family, where α is the vector of parameters. If, for every element $Y_\alpha \in \chi_\alpha$, the random variable $\frac{Y_\alpha - \mathbb{E}Y_\alpha}{\sqrt{Var(Y_\alpha)}}$ is independent of α and symmetric, then* $\max_a \mathbb{E}u\,(w(1+r)+a(Y-r))$ *is an increasing function of q_S.*

Normally distributed risk is a special case of this result.

Example 2. If Y_α is Normally distributed, then $\frac{Y_\alpha - \mathbb{E}Y_\alpha}{\sqrt{Var(Y_\alpha)}}$ is $N(0,1)$, which is independent of α and symmetric.

But so is the symmetric bivariate distribution.

Example 3. Consider the random variable X where $X = \alpha_1 + \sqrt{\alpha_2}$ w.p. 1/2, $X = \alpha_1 + \sqrt{\alpha_2}$ w.p. 1/2. Then $\frac{X_\alpha - \mathbb{E}X_\alpha}{\sqrt{Var(X_\alpha)}}$ is a bivariate random variable.

Indeed, we can construct many more probability spaces where the Sharpe ratio is a valid ranking measure.

Example 4. Suppose T is a t distribution with degree of freedom 4 and let $\chi_{\alpha_1,\alpha_2} = \{X : X = \alpha_1 + \alpha_2 * T\}$ be the set of all random variables (a distribution family) that can be written as linear function of T. Then $\forall Y_\alpha \in \chi_{\alpha_{1,\alpha_2}}$, the random variable $\frac{Y_\alpha - \mathbb{E}Y_\alpha}{\sqrt{Var(Y_\alpha)}} = \frac{T}{std(T)}$ is independent of parameter and symmetric. Hence, the Sharpe ratio properly ranks risky returns contained in the set χ_{α_1,α_2}.

These results demonstrate that Sharpe is potentially more robust than commonly understood. It holds even for non-Normal distributions without assuming quadratic utility. We explore the theoretical foundations of the Sharpe ratio in Section 4. However, when Sharpe ratio is not valid, it can "break, not bend." Consider the following example that comes from Hodges (1998).

Example 5. Consider two risky assets described by their risk net returns Y_1 and Y_2.

Probability	0.01	0.04	0.25	0.40	0.25	0.04	0.01
Excess Return Y_1	-25%	-15%	-5%	5%	15%	25%	35%
Excess Return Y_2	-25%	-15%	-5%	5%	15%	25%	45%

Clearly the first asset with return Y_1 is first-order stochastically dominated by the second asset with return Y_2. However, the Sharpe ratio for the first asset is 0.500 whereas the Sharpe ratio for the second asset is only 0.493.

Indeed, it is now generally understood that large Sharpe ratios can be produced simply by introducing options and other derivative contracts into the portfolio. We return to this topic in Section 6.

3 The Generalized Ranking Measure

We now derive our generalized ratio for ranking risks that is applicable to a broad admissible space.

3.1 The Regularity Condition

Using Taylor's theorem, we can rewrite the first-order condition of the investor's problem (1) as follows:

$$
\begin{aligned}
0 &= \mathbb{E}u'\left(w_r + a(Y-r)\right)(Y-r) \\
&= \mathbb{E}\left(\sum_{n=0}^{\infty} u^{(n+1)}(w_r)\frac{a^n(Y-r)^n}{n!}\right)(Y-r) \\
&= \sum_{n=0}^{\infty} \frac{u^{(n+1)}(w_r)\mathbb{E}(Y-r)^{n+1}}{n!}a^n \\
&= \sum_{n=1}^{\infty} \frac{u^{(n)}(w_r)\mathbb{E}(Y-r)^{n}}{(n-1)!}a^{n-1}
\end{aligned}
\tag{2}
$$

Definition 3. [The n-th t-moment] Define $t_n^Y \equiv \mathbb{E}(Y-r)^n$ as the *n-th translated moment* (the *n-th t-moment* for short) for the risky investment with return Y.

A closed-form solution of equation 2 is typically not available. (Section 5 provides some closed-form solutions for approximate measures.) Moreover, we can't let computers run indefinitely, and so we must truncate the expansion to a finite, even if large, number of terms, N. However, such a truncation, will typically produce *many* real and complex roots, even though the original infinite expansion in equation (2) has a single root by Lemma (1). Fortunately, the following lemma provides a mechanism for selecting the correct root in the N-term expansion, a result that likely has broader application outside of our current use.

Lemma 2. *Suppose real function $f(x) = 0$ has a unique real solution x_0. Denote the Maclaurin expansion of f to be $\sum_{n=0}^{\infty} c_n x^n$. Consider*

$$f_N(x) = \sum_{n=0}^{N} c_n x^n.$$

$f_N = 0$ has N solutions on the complex plane S_N. Denote the convergent radius for the series as λ. If $\lambda > |x_0|$, then: (i) the smallest absolute real root in S_N converges to x_0 as $N \to \infty$ and (ii) for any given bounded set, there is a finite value of N, call it $\bar{N}$, such that there is at most one real root in the set $\forall N > \bar{N}$.

Remark 2. The smallest absolute root does not necessarily converge monotonically (even in absolute value) as N grows. It is technically challenging to consider a "stopping rule" on the value of N. However, in practice, our computations suggest that the root does indeed converge after a reasonable value of N, especially after the only one real root emerges at large values of N. See Section 6.

Definition 4. [Regularity Condition] We will say that the utility-risk pair (u,Y) satisfies the *regularity condition* if the corresponding series of equation (2) satisfies the requirement $\lambda > |x_0|$ in Lemma 2. Denote $\mathscr{A}_{\mathscr{R}\mathscr{C}}$ as the admissible space of all the utility-risk pairs where the regularity condition holds.

Corollary 1. *The regularity condition trivially holds if the convergence radius is infinite (i.e., $\lambda = \infty$).*

Example 6. For CARA utility and any discrete distribution, the convergence radius is ∞.

3.2 HARA Utility

Consider the HARA utility function $u \in \mathscr{U}_H$ class where $u(w) = \frac{\rho}{1-\rho}\left(\frac{\lambda w}{\rho} + \phi\right)^{1-\rho}$. The HARA class is fairly broad and nests several other specifications as special cases, including decreasing, constant and increasing absolute risk aversion as well as decreasing, constant and increasing relative risk aversion. HARA is necessary and sufficient for asset demands to be linear in wealth (Merton, 1969), and the single-period ("myopic") problem with HARA utility has been the primary focus of the academic literature. Indeed, as Brandt (2010, P. 284) nicely puts it: "The myopic portfolio choice is an important special case for practitioners and academics alike. There are, to my knowledge, few financial institutions that implement multi-period investment strategies involving hedging demands."

Denote $\mathscr{U}_H^{\rho} \subset \mathscr{U}_H$ as the subset of HARA utility function where ρ is given. Then for $u \in \mathscr{U}_H^{\rho}$:

$$u^{(n)}(w) = \frac{\rho}{1-\rho}(1-\rho)(-\rho)\cdots(2-n-\rho)(\frac{\lambda}{\rho})^n\left(\frac{\lambda w}{\rho} + \phi\right)^{1-n-\rho}$$

From equation (2), we need to solve

$$\sum_{n=1}^{N} u^{(n)}(w_r)\frac{t_n^Y}{(n-1)!}a^{n-1} = 0 \tag{3}$$

as $N \to \infty$. With some direct substitutions, this series becomes:

$$\sum_{n=1}^{N} \frac{\rho}{1-\rho}(1-\rho)(-\rho)\cdots(2-n-\rho)(\frac{\lambda}{\rho})^n\left(\frac{\lambda w_r}{\rho} + \phi\right)^{1-n-\rho}\frac{t_n^Y}{(n-1)!}a^{n-1} = 0$$

or

$$\lambda\left(\frac{\lambda w_r}{\rho} + \phi\right)^{-\rho}\sum_{n=1}^{N}(\rho)\cdots(\rho+n-2)\frac{t_n^Y}{(n-1)!}\left(-\frac{\lambda}{\rho}\frac{a}{\frac{\lambda w_r}{\rho}+\phi}\right)^{n-1} = 0$$

Let

$$b_n = \begin{cases} 1, & n = 1 \\ (\rho)\cdots(\rho+n-2), & n \geq 2 \end{cases} \tag{4}$$

Also, let $z = -\frac{\lambda}{\rho}\frac{a}{\frac{\lambda w_r}{\rho}+\phi}$. With this change of variables, equation (3) can be rewritten as:

$$-\sum_{n=1}^{N}\frac{b_n t_n^Y}{(n-1)!}z^{n-1} = 0. \tag{5}$$

Definition 5. [Generalized Ranking Measure with HARA Utility] Let $z_{N,Y}$ denote the smallest absolute real root z that solves equation (5). The (N-th order) generalized ranking measure for HARA utility is:

$$q_H^N(t_n^Y, b_n) = -\sum_{n=1}^{N}\frac{b_n t_n^Y}{n!}z_{N,Y}^n. \tag{6}$$

where b_n is shown in equation (4) and t_n^Y is the n^{th} *t-moment* of the risky investment with return Y.

Notice that the root $z_{N,Y}$ is only a function of preferences b_n and the *t-moments* t_n^Y of the underlying risk distribution. In particular, $z_{N,Y}$ does not depend on the investor's wealth.

Theorem 2. *$q_H^N(t_n^Y, b_n)$ is a valid distribution-preference ranking measure sequence w.r.t. to the admissible space $\mathscr{A}_{\mathscr{H}} \equiv \left(\mathscr{U}_H^{\rho} \times \mathscr{Y}\right) \cap \mathscr{A}_{\mathscr{R}\mathscr{C}}$, where $\mathscr{Y}$ is the set of all random variables.*

In other words, q_H^N is a generalized ranking ratio that is valid for the HARA utility class under the regularity condition. A more general, but less relevant, measure can also be derived for preferences outside of the HARA class.[10] However, since the HARA utility class is already broad and the most relevant case in our (and the historical) context, we will simply refer to the distribution-preference ranking measure (6) as the *generalized risk ranking measure*, or the *generalized ratio* for brevity.

3.3 An Impossibility Theorem

Like the Sharpe ratio, the generalized ratio can rank risks without knowing the wealth of the underlying investors. The advantage of the generalized ranking ratio is that it can accommodate a broad admissible space. The disadvantage of the generalized ratio, however, is that it is a distribution-preference ranking measure that also requires knowledge of the investor's preferences, whereas the Sharpe ratio is a distribution-only ranking measure. The additional requirement of the generalized ratio, however, is not a feature of our particular construction. It is impossible to construct a distribution-only ranking measure that is generically valid for the HARA class.

Theorem 3. *There does not exist a distribution-only ranking measure for HARA utility if portfolio risk Y can be any random variable.*

Indeed, we can conclude that there is no generic distribution-only ranking measure when Y can take on any distribution, leading to the following impossibility theorem.

Corollary 2. *There does not exist a generic distribution-only ranking measure if portfolio risk Y can be any random variable.*

In other words, if we want to accommodate non-Normally distributed risk, a distribution-only ranking measure, like Sharpe, is not available across a wide range of investor preferences.

3.4 Two Special Cases

However, in two special cases of HARA utility, we can simplify things a bit more.

3.4.1 CARA Utility

In the case of constant absolute risk aversion (CARA), the value of $\phi = 1$, $\rho \to \infty$, we can show that it is equivalent to have $b_n = 1$. Hence, the corresponding value of $z_{N,Y}$ is only a function of the *t-moments* of the underlying risk, due to the absence of the income effect.

Corollary 3. *For CARA utility, if the regularity condition holds, a distribution-only ranking measure exists and takes the form $q_{CARA}(t_n^Y) = -\sum_{n=1}^{N} \frac{t_n^Y}{n!} z_{N,Y}^n$.*

[10] In particular, we can drop the assumption of HARA utility if we are comfortable with a potentially less powerful *distribution-preference-wealth* ranking measure that also requires knowing the investor's wealth. In particular, the ranking measure would take the form $q_H^N(t_n^Y, u, w_r) = -\sum_{n=1}^{N} \frac{u^{(n)}(w_r) t_n^Y}{n!} z_{N,Y}^n$, where z solves a finite N version of equation equation (2). In this case, relative to the original investor problem (1), we don't need to know the underlying risk distribution; we only need to estimate its *t-moments*. Still, requiring information about the investor's wealth is much more restrictive for investment fund managers. Moreover, asset demand is no longer linear in wealth.

In other words, the valid ranking measure for CARA utility only requires only a characterization of the shock distribution, just like the Sharpe ratio. Unlike the Sharpe ratio, however, this measure is valid for non-Normally distributed risk (if preferences are CARA). Of course, while CARA utility is commonly used for theoretical analysis, its applicability to actual investor problems is quite limited.

3.4.2 CRRA Utility

As with the Sharpe ratio, the generalized ratio only ranks various *risky* portfolios in search for the best one. Neither the Sharpe ratio or the generalized ratio determines the optimal split between the best risky asset and the risk-free asset. That division must be determined in a second stage using consumer preferences and wealth. In the finance industry, that's where the division between the investment manager and the personal financial advisor comes into play. The investment manager searches for the best stock portfolio that produces the largest ranking measure, without specific knowledge of the underlying investors. The financial planner then gets to know the risk tolerance of each client to determine the optimal split between stocks and bonds.

However, we can collapse both of these steps into a single step in the case of constant relative risk aversion (CRRA) utility where $\phi = 0$, $\rho > 0$, and $\lambda = \rho$. The CRRA form is commonly used to model investor preferences since the level of risk aversion scales with investor wealth. To be sure, the CRRA ranking function is still a *distribution-preference* ranking measure, as in the more general HARA case. But, we obtain a nice simplification.

Remark 3. For CRRA utility, the quantity $-z(1+r)$ is equal to a/w, the percentage of wealth w that is invested into the risky asset.

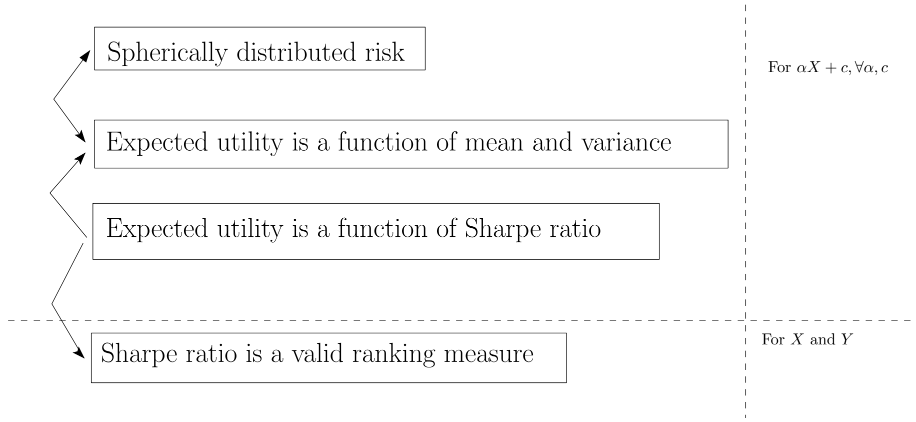
In other words, upon picking the best risky investment Y from the admissible space $\mathscr{A}_{\mathscr{H}}$, we can also immediately determine the share of wealth to be placed into this risky investment (versus bonds). However, the quantity $-z(1+r)$ itself is not a valid ranking measure since the generalized ranking measure under CRRA is not a monotone transformation of $-z(1+r)$.[11]

3.5 Extension to Multiple Asset Classes

The set $\mathscr{Y}$ of random variables in Theorem 2 potentially includes random variables drawn from different distributions. The generalized ratio, therefore, allows one to consider *composite risks*, that is, multi-asset class portfolios. Examples include portfolios with derivatives, thinly traded securities, and corporate bonds. The generalized ratio, therefore, can be used as the foundation for multi-asset class portfolio optimization due to its ability to correctly pairwise rank composite risks following different distributions. For example, one can pairwise rank random return Y_1, representing a portfolio without derivatives, against random return Y_2, representing a portfolio with derivatives following a different distribution. This flexibility is more powerful than simply assuming that all potential portfolio combinations follow the same probability distribution form, even if that distribution is more extensible than Normality. The only additional practical step required for full portfolio optimization is to combine the ranking function q_H with a globally stable optimization routine that searches over the space of potential composite assets contained within $\mathscr{A}_{\mathscr{H}}$. We provide examples in Section 6.

[11]For example, consider two risky asset payoffs, Y and tY, where t is a positive constant. The generalized ranking measure produces identical values since the investor should be indifferent between the two risky assets. However, the percent invested into each asset will differ.

Figure 1: Relationship Between Spherically Distributed Risks and the Sharpe Ratio



4 Theoretical Foundations of the Sharpe Ratio

Despite its widespread use, the Sharpe ratio is still a bit of a mystery because it holds between seemingly unrelated admissible spaces. It correctly ranks risks that are Normally distributed for those types of preferences consistent with an unbounded distribution. It is also correctly ranks non-Normally distributed risks if preferences are (albeit, unrealistically) quadratic. Furthermore, as we showed in Section 2, the Sharpe ratio is valid under more general admissible spaces with non-Normal risks and non-quadratic preferences. This section explores the foundations of the popular Sharpe ratio in more detail. We are able to obtain key insights by shrinking the investment time horizon to near zero and then, under certain conditions, map back to the original investor problem considered in Sections 2 and 3. In the process, we also extend the classic Kraus-Litzenberger (1976) result, demonstrating the preference of investors for skewed returns, to infinite (adjusted) cumulants.

4.1 An Important Aside: Understanding Spherically Distributed Risks

At a first glance, our claim that it is challenging to derive a necessary condition for the Sharpe ratio to be a valid ranking measure might appear to be at odds with Chamberlain (1983). Chamberlain nicely demonstrates that the distribution of every portfolio can be fully described by its mean and variance if and only if its random return is a linear transformation of a Spherically distributed random vector.

In practice, this important result has been incorrectly interpreted by the industry to imply that a Spherically distributed risk is a necessary and sufficient condition for the Sharpe ratio to be a valid ranking measure. That interpretation, however, makes two mistakes, as we illustrate graphically in Figure 1. First, Chamberlain's result simply says if the underlying risk is Spherically distributed, then there exists *some* function of its mean and variance that correctly summarizes the investor's expected utility. That function does not have to take the form of the *Sharpe* ratio, however. Second, and even more fundamentally, Chamberlain's isomorphic mapping is operating on the space of functions that are a linear transformation of the Spherically distributed random variable X. That domain is different, however, than that of the Sharpe ratio ranking measure which is comparing two risks, X and Y. The following example shows that the Sharpe ratio can incorrectly rank two risks that are both Spherically distributed.

Example 7. Suppose an investor's utility function takes the CARA form, $u(w) = -e^{-w}$. Risky investments X and Y take following forms:

$$X = \begin{cases} x_1 & \text{w.p } 0.5 \\ x_2 & \text{w.p } 0.5 \end{cases}$$

$$Y = \begin{cases} y_1 & \text{w.p } \frac{1}{3} \\ y_2 & \text{w.p } \frac{1}{3} \\ 2y_2 - y_1 & \text{w.p } \frac{1}{3} \end{cases}$$

where $x_1 < 0 < x_2, y_1 < y_2, y_1 < 0, 2y_2 > y_1$. Thus, both X and Y are Spherically distributed. Assume the risk free rate is 0. As before, the standard investor problem is to solve for the amount of his wealth to be invested into each risk, respectively:

$$a_X = arg \max_a \mathbb{E}u(w + aX)$$

$$a_Y = arg \max_a \mathbb{E}u(w + aY)$$

Since CARA doesn't have income effect, we can set wealth $w = 0$ without any loss in generality. Hence, we can solve for a_X and a_Y analytically: $a_X = \frac{\ln(-\frac{x_2}{x_1})}{x_2 - x_1}$ and $a_Y = -\frac{\ln(\frac{-y_2 + \sqrt{y_2^2 - 4y_1(2y_2 - y_1)}}{2(2y_1 - y_1)})}{y_2 - y_1}$. So the maximum expected utility that the investor can achieve if he chooses X is

$$-exp(-a_X x_1)/2 - exp(-a_X x_2)/2.$$

If, instead, the investor chooses Y, his maximum expected utility is

$$-exp(-a_Y y_1)/3 - exp(-a_Y y_2)/3 - exp(-a_Y y_3)/3.$$

Now set $x_1 = -0.05$, $x_2 = 0.35$, $y_1 = -0.0473$, and $y_2 = 0.0752$. Then, the Sharpe ratio for investment X is $\frac{0.15}{0.2} = 0.75$ while the Sharpe ratio for Y is $\frac{0.0752}{0.1} = 0.752$. But, the maximized expected utilities for X and Y are -0.7288 and -0.7359, respectively. In other word, Spherically distributed risk Y has a higher Sharpe Ratio than Spherically distributed X, but the investor actually prefers X over Y.

Remark 4. Example 7 uses CARA utility with a discrete distribution. Hence, by Example 6, its convergence radius is ∞ and, by Corollary 1, its regularity condition holds. The generalized ranking measure, therefore, correctly ranks X over Y. Moreover, by Corollary 3, it is a distribution-only ranking measure.

4.2 Adjusted Cumulants

This subsection develops the preliminary mathematical concepts required for analyzing risk ranking as the investment horizon approaches zero.

4.2.1 Definitions

Definition 6. [Infinitely Divisible] For a given probability space, we say that random variable Y has an infinitely divisible distribution, if for each positive integer T, there is a sequence of i.i.d. random variables $X_{T,1}$, $X_{T,2}$,...,$X_{T,T}$ such that

$$Y \stackrel{d}{=} X_{T,1} + X_{T,2} + \cdots + X_{T,T},$$

where the symbol $\stackrel{d}{=}$ denotes equality in distribution. We say Y has the "infinitely divisibility property."

We can think of a single unit of time as being divided into T subintervals of equal length of time, Δt, i.e., $\Delta t = \frac{1}{T}$. Each variable $X_{T,i}$ then represents the return during the i-th subinterval. For notational simplicity, since the $X_{T,1}, X_{T,2}, \cdots, X_{T,T}$ subintervals of risk Y are i.i.d., we drop the subscripts and simply express each subinterval as X.

Definition 7. [Adjusted Cumulant] Suppose Y has an infinitely divisible distribution and let

$$\varepsilon_T = \frac{X - \frac{\mu}{T}}{\sigma\sqrt{\frac{1}{T}}} = \frac{X - \mu\Delta t}{\sigma\sqrt{\Delta t}},$$

where μ and σ are the mean and standard deviation of Y. In the context of Δt , we write $\varepsilon = \frac{X-\mu\Delta t}{\sigma\sqrt{\Delta t}}$. Define Y's n-th's *adjusted cumulant* as

$$\nu_n = \nu_n(Y) = \lim_{T\to\infty} \frac{\mathbb{E}\varepsilon_T^n}{\left(\frac{1}{T}\right)^{-\frac{n-2}{2}}} = \lim_{\Delta t\to 0} \frac{\mathbb{E}\varepsilon^n}{(\Delta t)^{-\frac{n-2}{2}}}, \forall n \geq 2.$$

Lemma 3. *The more traditionally defined (that is, non-adjusted) n-th cumulant is equal to $\nu_n\sigma^n$, $\forall n \geq 2$.*

The adjusted cumulant concept is easier to interpret than the traditional cumulant of a distribution. In particular, ν_3 corresponds to a random variable Y's skewness while ν_4 is its excess kurtosis, etc. Moreover, like the traditional cumulant, if Y is Normally distributed then $\nu_n = 0$, $\forall n \geq 3$.

4.2.2 Calculating Adjusted Cumulants

For random variable Y, denote μ_1 as the mean and for $n \geq 2$,

$$\mu_n = \mathbb{E}(Y - \mathbb{E}Y)^n \text{ and } \xi_n = \frac{\mu_n}{\mu_2^{\frac{n}{2}}}$$

So, ξ_3 represents the skewness and ξ_4 the kurtosis. We can then calculate adjusted cumulants either by induction or using the known distribution's moment generating function. Each approach has its relative strengths. First, consider the induction approach.

Theorem 4. *For any integer $n \geq 4$,*

$$\nu_n = \xi_n - \sum_{\substack{n = i_1 + i_2 + \cdots + i_k \\ i_1 \geq i_2 \cdots \geq i_k \geq 2}} \frac{\binom{n}{i_1}\binom{n-i_1}{i_2}\cdots\binom{n-i_1-\cdots-i_{k-1}}{i_k}}{k!} \nu_{i_1}\cdots\nu_{i_k}.$$

The key advantage of the inductive approach is that adjusted cumulants can be easily calculated using actual data where the functional form of the risk distribution is not known. While that precise purpose is not the focus of this section, it does make the point that our derivations are not reliant on being able to calculate the cumulants of a known distribution. Theorem 4 likely has value outside of the current study.

Remark 5. By Remark 3, the *standard* cumulant of a distribution can, therefore, also be computed inductively using Theorem 4.

Another way to calculated adjusted cumulants is by exploiting the fact that an infinitely divisible distribution corresponds to a Levy process. Suppose X_t is Levy Process and $Y = X_1$. Then Y is infinitely divisible distribution with $X = X_{\frac{1}{T}}$. By Levy-Khinchine representation, we have

$$\mathbb{E}e^{i\theta X_t} = exp\left(bit\theta - \frac{1}{2}\sigma_0^2 t\theta^2 + t\int_{\mathbb{R}\setminus\{0\}}(e^{i\theta x} - 1 - i\theta x I_{|x|<1})W(dx)\right)$$

where $b \in \mathbb{R}$, I is the indicator function and i is the imaginary unit and θ is the parameter for the characteristic function. The Levy measure W must be such that

$$\int_{\mathbb{R}\setminus\{0\}} min\{x^2, 1\}W(dx) < \infty$$

Denote

$$\phi(\theta,t) \equiv bit\theta - \frac{1}{2}\sigma_0^2 t\theta^2 + t\int_{\mathbb{R}\setminus\{0\}}(e^{i\theta x} - 1 - i\theta x I_{|x|<1})W(dx)$$

and

$$\psi(\theta,t) \equiv bt\theta + \frac{1}{2}\sigma_0^2 t\theta^2 + t\int_{\mathbb{R}\setminus\{0\}}(e^{\theta x} - 1 - \theta x I_{|x|<1})W(dx)$$

i.e., $\phi(\theta,t) = \psi(i\theta,t)$ where $e^{\psi(\theta,t)}$ is the moment generating function of X_t.

Theorem 5. *Suppose $e^{\psi(\theta,t)}$ is the moment generating function of the Levy Process X_t. Let $Y = X_1$ and let σ denote the standard deviation of Y. Then:*

$$\nu_n(Y) = \frac{\frac{\partial^n \psi(\theta,1)}{\partial\theta^n}|_{\theta=0}}{\sigma^n}, \forall n \geq 2.$$

4.3 Ranking Measure with Short Trading Times

We are now ready to consider the investor problem as the time horizon approaches zero. For the risky return Y with an infinitely divisible distribution, consider the Δt period investor problem.

$$\begin{aligned} &\max_a \mathbb{E}u(w(1+r\Delta t) + a(X - r\Delta t)) \qquad (7) \\ &= \max_a \sum_{n=0}^{+\infty} u^{(n)}(w(1+r\Delta t))\frac{a^n}{n!}\mathbb{E}(X - r\Delta t)^n \\ &= \max_a \sum_{n=0}^{+\infty} u^{(n)}(w(1+r\Delta t))\frac{a^n}{n!}\mathbb{E}(\mu\Delta t + \sigma\sqrt{\Delta t}\varepsilon - r\Delta t)^n \end{aligned}$$

By definition of adjusted cumulants, the leading term of $\mathbb{E}(\mu\Delta t + \sigma\sqrt{\Delta t}\varepsilon - r\Delta t)^n$ is $\sigma^n(\Delta t)^{\frac{n}{2}}\mathbb{E}\varepsilon^n$, which is of order $\sigma^n\nu_n\Delta t$ for $n \geq 2$, and it is $(\mu - r)\Delta t$ when $n = 1$. Denote $\nu_1 = \frac{\mu - r}{\sigma}$. Then $\mathbb{E}(\mu\Delta t + \sigma\sqrt{\Delta t}\varepsilon - r\Delta t)^n \sim \sigma^n\nu_n\Delta t$ for any $n \geq 1$. Denote $w_r = w(1+r\Delta t)$. So

$$\begin{aligned} &\max_a \sum_{n=0}^{+\infty} u^{(n)}(w(1+r\Delta t))\frac{a^n}{n!}\mathbb{E}(\mu\Delta t + \sigma\sqrt{\Delta t}\varepsilon - r\Delta t)^n \\ &= \max_a \left(u(w_r) + \sum_{n=1}^{+\infty} u^{(n)}(w_r)\frac{\sigma^n\nu_n\Delta t}{n!}a^n + o(\Delta t)\right) \\ &= \max_a \left(u(w_r) + \left(\sum_{n=1}^{+\infty} u^{(n)}(w_r)\frac{\nu_n}{n!}(\sigma a)^n\right)\Delta t + o(\Delta t)\right). \end{aligned}$$

Theorem 6. *As $\Delta t \to 0$, the maximization problem (7) is*

$$\sum_{n=1}^{+\infty} u^{(n)}(w_r) \frac{\nu_n}{(n-1)!} (\sigma a)^{n-1} = 0, \tag{8}$$

assuming that this series converges.

Now, consider the *finite* series:

$$\sum_{n=1}^{N} u^{(n)}(w_r) \frac{\nu_n}{(n-1)!} (\sigma a)^{n-1} = 0 \tag{9}$$

and let a_N^* equal the smallest absolute real root that solves equation (9). By Lemma 2, this value will converge to the root of series (8) for a large enough value of N, if the series' regularity condition holds. Inserting a_N^* into the investor problem:

$$\begin{aligned}
&\max_a \mathbb{E}u(w(1+r\Delta t) + a(X - r\Delta t)) \\
&= \mathbb{E}u(w(1+r\Delta t) + a_N^*(X - r\Delta t)) \\
&= u(w_r) + \left(\sum_{n=1}^{+\infty} u^{(n)}(w_r) \frac{\nu_n}{n!} (\sigma a_N^*)^n \right) \Delta t + o(\Delta t) \\
&\approx u\left(w_r + \frac{\left(\sum_{n=1}^{+\infty} u^{(n)}(w_r) \frac{\nu_n}{n!} (\sigma a_N^*)^n \right) \Delta t}{u'(w_r)} \right) \\
&= u\left(w_r + \left(\sum_{n=1}^{+\infty} \frac{u^{(n)}(w_r)}{u'(w_r)} \frac{\nu_n}{n!} (\sigma a_N^*)^n \right) \Delta t \right)
\end{aligned}$$

Definition 8. [Adjusted Cumulant Ranking Measure] Now define the *N-th order adjusted cumulant ranking* measure as

$$\sum_{n=1}^{N} \frac{u^{(n)}(w_r)}{u'(w_r)} \frac{\nu_n}{n!} (\sigma a_N^*)^n \tag{10}$$

As N approaches to infinity, we simply call it the *adjusted cumulant ranking measure.*

Example 8. If utility takes the CARA form, the adjusted cumulant ranking measure takes the simple form $\sum_{n=1}^{N} \frac{\nu_n}{n!} (\sigma a_N)^n$.[12]

Of course, the adjusted cumulant ranking measure, which is defined over the $\Delta t \to 0$ period, is not practical unless it gives us some insight into the discrete time length T investor problem considered in Sections 2 and 3. The following definition and results allow us to make that connection.

[12] Dávila (2010) derives a similar expansion for CARA utility for $N = \infty$. Since he did not have the concept of Lemma (2), he did not solve the finite N case in order to derive a ranking measure for CARA with general cumulants. Instead, he considers some specific distributions that are assumed to be known ex-ante. To be sure, CARA utility is restrictive for practical purposes. Moreover, assuming that each risk comes from the same known distribution prohibits making pairwise comparisons across different risk distributions, which, for example, effectively rules out multi-asset class optimization. Nonetheless, his nice paper appears to have made the most headway in the first line of literature referenced in Section 1.

Definition 9. [Scalable] We say that utility function u is *scalable* with respect to the random variable space χ, a subset of all infinite divisible distributions, if the following equivalence holds

$$(Y,r) \geq_u (Y',r) \Leftrightarrow (X,r\Delta t) \geq_u (X',r\Delta t)$$

where $Y,Y' \in \chi$ and X and X' are any subinterval of Y and Y' of length Δt. In other words, if u is scalable, an investor prefers Y over Y' if and only if he prefers X over X' in the Δt time period.

Theorem 7. *If u is scalable with respect to random variable space χ, then the adjusted cumulant ranking measure is also a valid ranking measure for the original discrete-time investor problem (1) with respect to $u \times \chi$.*

Lemma 4. *CARA utility and Quadratic utility are both scalable with respect to all infinite divisible distributions. HARA utility is scalable with respect to all Poisson distributions. Any utility function is scalable with respect to a Normal risk distribution.*

Example 9. Suppose utility is CARA and the underlying portfolio risk distributions are infinitely divisible that satisfy the regularity condition. Then, the adjusted cumulant ranking measure is a valid distribution-only ranking measure, that is,

$$\sum_{n=1}^{+\infty} \frac{\nu_n^{Y_1}}{n!}(\sigma_1 a_{N,Y_1})^n > \sum_{n=1}^{+\infty} \frac{\nu_n^{Y_2}}{n!}(\sigma_2 a_{N,Y_2})^n \Longleftrightarrow (Y_1,r) \geq_u (Y_2,r)$$

where σ_i is the standard deviation of Y_i, $i = 1,2$. In other words, with CARA utility, ranking the Δt subinterval problem with the adjusted cumulant measure correctly ranks the original investor problem where the investment problem is made over the discrete time length T. Intuitively, the absence of wealth effects with CARA utility means that there is no need for rebalancing.

Of course, many utility functions are not scalable with common probability distributions. Moreover, in practice, the probability distribution is often unknown. Hence, the generalized ratio derived in Section 3 must be used to rank risky returns for the original discrete time length T investor problem (1), rather than the adjusted cumulant ranking measure.

Still, the $\Delta t \to 0$ analysis developed in this section allows us to gain deeper insights that carry over to the discrete time length T problem *if* the utility function is scalable with respect to the underlying risk distribution. The most relevant cases for analyzing the Sharpe ratio are indeed scalable. By Lemma 4, these cases include quadratic preferences combined with any infinitely divisible distribution as well as any utility function combined with Normally distributed risk.[13]

4.4 Understanding the Sharpe Ratio

We are now well positioned to develop a deeper understanding of the Sharpe Ratio. Suppose that, for all $n \geq 3$, $\nu_n = 0$ or $u^{(n)} = 0$. The first-order condition (8) is then

$$\sum_{n=1}^{2} u^{(n)}(w_r)\frac{\nu_n}{(n-1)!}(\sigma a)^{n-1} = 0 \tag{11}$$

[13]Of course, as noted earlier, the investor problem (1) should be well-defined. For example, the Inada condition ($u'(w \to 0) = \infty$) associated with CRRA utility implies that the demand for a Normally distributed risky asset would always be zero due to unlimited liability.

and the 2nd-order adjusted cumulant ranking measure implies:

$$\sigma a^* = -\frac{u'(w_r)\nu_1}{u''(w_r)\nu_2} = -\frac{u'(w_r)}{u''(w_r)}\frac{\mu - r}{\sigma}$$

The investor's indirect utility is then given by:

$$\begin{aligned}
&\max_a \mathbb{E}u(w_r + a(X-r)) \\
&= u(w_r) + \left(\sum_{n=1}^{2} u^{(n)}(w_r)\frac{\nu_n}{n!}(\sigma a^*)^n\right)\Delta t + o(\Delta t) \\
&= u(w_r) + \left(\sum_{n=1}^{2} u^{(n)}(w_r)\frac{\nu_n}{n!}\left(-\frac{u'(w_r)}{u''(w_r)}\frac{\mu - r}{\sigma}\right)^n\right)\Delta t + o(\Delta t) \\
&= u(w_r) + u'(w_r)\left(-\frac{u'(w_r)}{u''(w_r)}(\frac{\mu - r}{\sigma})^2 + \frac{1}{2}\frac{u'(w_r)}{u''(w_r)}(\frac{\mu - r}{\sigma})^2\right)\Delta t + o(\Delta t) \\
&= u(w_r) - u'(w_r)\left(\frac{1}{2}\frac{u'(w_r)}{u''(w_r)}(\frac{\mu - r}{\sigma})^2\right)\Delta t + o(\Delta t) \\
&\approx u\left(w_r - \frac{1}{2}\frac{u'(w_r)}{u''(w_r)}(\frac{\mu - r}{\sigma})^2\Delta t\right) \\
&= u\left(w_r - \frac{1}{2}\frac{u'(w_r)}{u''(w_r)}\left(\frac{\mu\Delta t - r\Delta t}{\sigma\sqrt{\Delta t}}\right)^2\right)
\end{aligned}$$

Notice that the Sharpe Ratio of X is $\frac{\mu\Delta t - r\Delta t}{\sigma\sqrt{\Delta t}}$. Hence, the Sharpe ratio is a valid ranking measure.[14]

We can, therefore, now see why the Sharpe ratio works for Normal risk distributions or quadratic utility. The reason stems directly from the multiplicative pairing of the Nth-order marginal utility $u^{(n)}$ and the Nth-order adjusted cumulant ν_n within each separable term inside of the investor's first-order condition (9). Anything that "zeros out" either one of these 3rd- and higher-order terms is sufficient to make the Sharpe ratio valid. In particular, if risk Y is Normally distributed, then $\nu_n = 0$, $n \geq 3$. If u takes the quadratic form then $u^{(n)} = 0$, $n \geq 3$.

It is also clear why Normally distributed risk or quadratic utility are not *necessary* conditions either for Sharpe to be valid. Indeed, it is possible that equation (11) emerges if, for example, $\nu_n = 0$ for odd values of n and $u^{(n)} = 0$ for even values, or some other combination.

In fact, in turns out that even equation (11) is not even a necessary condition for the Sharpe ratio to be valid. In fact, some of the examples we provided in Section 2, where Sharpe is a correct ranking measure, do *not* produce equation (11). In other words, an even more general sufficient condition for Sharpe exists. We provide one in the following Theorem, which nests equation (11) as a special case.

Theorem 8. *For given utility function u, suppose that the corresponding risk space χ_u is not empty. Then, the Sharpe ratio is a valid ranking measure on set χ_u if the higher-order adjusted cumulants of all risks in set χ_u are equal to each other (i.e. $\nu_k(Y) = \nu_k(Y'), \forall Y, Y' \in \chi_u, k \geq 3$), with the odd-numbered cumulants equal to zero (i.e., $\nu_k(Y) = \nu_k(Y') = 0, \forall Y, Y' \in \chi_u, k = 3, 5, 7, \ldots$).*

[14] Technically, the ranking measure is the square of the Sharpe Ratio, which implies the Sharpe Ratio when the expected equity premium, $\mu - r$, is positive. For brevity, we won't continue to make this distinction in the discussion below under the assumption that, in equilibrium, risky securities must pay a risk premium.

Example 10. Suppose T is a given symmetric infinitely divisible distribution. Construct a new distribution family of the form $\chi_{\alpha_1,\alpha_2} = \{X : X = \alpha_1 + \alpha_2 * T\}$. Then, the adjusted cumulants are $\nu_{2k}(X) = \nu_{2k}(T)$ and $\nu_{2k+1}(X) = \nu_{2k+1}(T) = 0$, $\forall k \geq 1$. Therefore, the Sharpe ratio is a valid ranking measure on the set χ_{α_1,α_2}.

4.5 Generalization of the Skewness Preference

Using a three-moment distribution, Kraus and Litzenberger (1976) well-cited paper demonstrates that investors with cubic utility prefer skewness in their returns.[15] More recently, Peirro and Mosevich (2011) demonstrate that, in the special case of CARA utility, investors dislike kurtosis as well. Dávila (2010) considers higher-order terms for the CARA utility case.

Most interesting utility functions have infinite non-zero higher-order terms. Moreover, on the risk side, the case of Normally distributed risk $(\nu_k = 0, k > 2)$ turns out to be extremely special. In particular, there does not exist a probability distribution that can be characterized by just adding a finite number of additional higher-order cumulants in order to expand on mean-variance analysis.

Lemma 5. *There does not exist a random probability distribution for which $\nu_m = \nu_{m+1} = ... = 0$ for some $m > 3$, with the lower-order adjusted cumulants (orders 3 to $m-1$) being nonzero.*

Hence, it is interesting to consider high-order terms as well. The following theorem generalizes the Kraus-Litzenberger result.

Theorem 9. *If $\mu > r$, the adjusted cumulant risk measure (equation (10)) is increasing with respect to odd adjusted cumulants $\nu_3, \nu_5, \cdots$ and decreasing with respect to even adjusted cumulants $\nu_4, \nu_6, \cdots$. If $\mu < r$, the adjusted cumulant risk measure is decreasing with respect to odd adjusted cumulant $\nu_3, \nu_5, \cdots$ and increasing with respect to even adjusted cumulant $\nu_4, \nu_6, \cdots$.*

Theorem 9, therefore, generalizes the previous results either in the number of higher-order terms or in the allowable admissible utility-probability space. For this result to be applicable to the discrete time length T problem considered in Section 3, we only require that the utility function is scalable with respect to the underlying risk distribution.

Corollary 4. *Suppose $\mu > r$, the investor prefers high skewness and low kurtosis.*

5 An Approximation Formula

Historically, there was a significant debate whether Markowitz's (1952) mean-variance foundation, which is central to the Sharpe ratio, was a useful approximation even when the underlying risk was not Normally distributed. On one hand, Borch (1969) and Feldstein (1969) argued that mean-variance framework was not robust. On the other hand, Tobin (1969), Markowitz (1959), and Samuelson (1970) argued that mean-variance was, in fact, a good approximation. Samuelson (1970, P. 542) concludes his article as follows: "Notice how higher-than second moments do improve the solution. But it also needs emphasizing that ... when 'risk is quite limited,' the mean-variance result is a very good approximation. When the heat of the controversy dissipates, that I think will be generally agreed on." Using a market data on a broad stock index, Levy and Markowitz (1979) and Kroll, Levy and Markowitz (1984) were

[15] See also Kane (1982).

influential in suggesting that the investor's expected utility problem could be well approximated by the mean and variance.

Of course, this debate took place before newer asset classes, beyond the broad stock market index, became more common, and certainly before investment managers started using derivatives to manage risks which produce non-Normal risk distributions. Without Lemma 2 from Section 3, allowing for additional higher-order terms for investment ranking often historically took the form of adding functional approximations of higher-order terms to the mean-variance model (as oppose to a convergence to the exact risk ranking as N grows).

The question we address in this section and Section 6 is whether those approximations are essentially "good enough," and, hence, does our generalized ratio really add that much value. To give the approximation approach its best chance, we extend it to an infinite number of higher moment terms. As a preview, the computations in Section 6 show that these approximations still often fail to correctly rank portfolios, are unstable, and can even diverge. To be clear, our intent is not to critique previous attempts to extend the Sharpe ratio to allow for additional moments using approximation methods. Rather, our motivation is to demonstrate that the generalized ratio, derived in Section 3, provides a lot more ranking power than can be achieved simply by adding functional approximations of higher-order terms to the mean-variance framework.

Although it is impossible to get a closed-form formula for the investor's first-order condition (9), we can solve for a linearized closed-form solution. Starting with equation (9),

$$\sum_{n=1}^{N} \frac{u^{(n)}(w_r)}{u'(w_r)} \frac{\nu_n}{(n-1)!} (\sigma a)^{n-1} = 0,$$

we can rearrange to solve,

$$\frac{\mu - r}{\sigma} + \frac{u''(w_r)}{u'(w_r)}(\sigma a) + \sum_{n=3}^{N} \frac{u^{(n)}(w_r)}{u'(w_r)} \frac{\nu_n}{(n-1)!} (\sigma a)^{n-1} = 0,$$

for its root σa^*, which depends only on the coefficients of the polynomial. Now denote

$$\sigma a^* = -\frac{u'(w_r)}{u''(w_r)} \frac{\mu - r}{\sigma} + g\left(\frac{u^{(n)}(w_r)}{u'(w_r)} \nu_n; 1 \le n \le N \right)$$

The most straightforward approximation of g as $\nu_n \to 0$ for all $3 \le n \le N$ is a linear function of $\{\nu_n; 3 \le n \le N\}$. Suppose

$$g\left(\frac{u^{(n)}(w_r)}{u'(w_r)} \nu_n; 1 \le n \le N \right) \approx \sum_{n=3}^{N} c_n \frac{u^{(n)}(w_r)}{u'(w_r)} \nu_n$$

i.e.

$$\sigma a^* \approx -\frac{u'(w_r)}{u''(w_r)} \frac{\mu - r}{\sigma} + \sum_{n=3}^{N} c_n \frac{u^{(n)}(w_r)}{u'(w_r)} \nu_n$$

Define $p_n = \frac{u^{(n)}(w_r)}{u'(w_r)}$. Then:

$$\sigma a^* \approx -\frac{1}{p_2} \nu_1 + \sum_{n=3}^{N} \frac{(-\frac{1}{p_2})^n \nu_1^{n-1}}{(n-1)!} p_n \nu_n. \tag{12}$$

Theorem 10. *If equation (12) holds with equality, then the investor's expected utility is increasing in the value of*

$$-\frac{\nu_1^2}{2p_2} + \sum_{n=3}^{N} \frac{p_n \nu_n}{n!} (-\frac{\nu_1}{p_2})^n \tag{13}$$

which we will call the "approximate ranking measure."

Remark 6. For HARA utility, the approximate ranking measure becomes:

$$-\frac{1}{p_2}(\frac{\nu_1^2}{2} + \sum_{n=3}^{N} \frac{b_n \nu_1^n \nu_n (-1)^{n-1}}{n!}) \tag{14}$$

Like the generalized ratio derived in Section 3, the approximate measure (14) has a nice property in that it converges to the Sharpe ratio as the higher-order adjusted cumulants $(\nu_n = 0, n > 2)$. It also nests some other extensions considered in the literature:

- When $N = 2$, we have $-\frac{1}{p_2}(\frac{SR^2}{2})$, where SR denotes the Sharpe Ratio, corresponding to the mean-variance framework tested empirically by Levy and Markowitz (1979).
- When $N = 3$, we have $\frac{-1}{p_2}\left(\frac{SR^2}{2} + \frac{b_3}{6} SR^3 (\nu_3 (\Delta t)^{-/2})\right)$. Notice that $\nu_3(\Delta t)^{-/2}$ is the skewness of X. This formula matches the extension of the mean-variance framework by Zakamouline and Koekebakker (2008) to include skewness.
- When $N = 4$, we have $\frac{-1}{p_2}\left(\frac{SR^2}{2} + \frac{b_3}{6} SR^3 * Skew - \frac{b_4}{24} SR^4 * (Kurt - 3)\right)$, where $Skew$ corresponds to the skewness. Di Pierro and Mosevich (2011) derive a similar formula, but specialized to CARA investors and no risk-free asset, allowing them to avoid explicit linearization and instead group the approximation error into a single term.

6 Additional Applications

6.1 Hodges' (1998) Paradox

Let's now return to Hodges' paradox that was considered earlier in Example 5.

Probability	0.01	0.04	0.25	0.40	0.25	0.04	0.01
Excess Return of Asset A	−25%	−15%	−5%	5%	15%	25%	35%
Excess Return of Asset B	−25%	−15%	−5%	5%	15%	25%	45%

As noted in Section 2, Asset B clearly first-order stochastically dominates Asset A. However, Asset A has a Sharpe ratio of 0.500, whereas Asset B has a Sharpe ratio of 0.493. We now use our generalized ratio to re-evaluate this paradox. Even with the distribution-only CARA version of our ranking function, the generalized ratio is able to correctly rank Asset B greater than Asset A at a value of $N \geq 5$ or more adjusted cumulants (Table (1)). Our approximation formula also seems to work reasonably well in this case (Table (2)).

Table 1: Hodges' Paradox – The Generalized Ratio

N	3	4	5	10
Asset A	NaN	0.1150	0.1172	0.1166
Asset B	NaN	0.1140	0.1190	0.1173

Explanation: Ranking measures for the distribution-only CARA version of the generalized ratio for Hodge's example, where N is the largest adjusted cumulant used in the shown calculation.

Table 2: Hodge's Paradox – Approximation Formula

N	3	4	5	10
Asset A	0.1237	0.1227	0.1227	0.1227
Asset B	0.1263	0.1228	0.1239	0.1236

Explanation: Ranking measures for the distribution-only CARA version of the approximate ratio for Hodge's example, where N is the largest adjusted cumulant used in the shown calculation.

6.2 Single Fund Asset Allocation

As noted in Section 3.4.2, in the case of CRRA utility, we can solve the "investment manager" and "financial planner" problem simultaneously as part of generating the ranking index. Combining the generalized ratio with a globally stable optimizer,[16] we, therefore, calculate the *optimal asset allocation* into the S&P500, based on monthly returns from January 1950 to June 2012, versus a risk-free bond paying an annual interest rate $r_f = 5\%$. In particular, for each level of risk aversion, a global optimizer finds the stock allocation that maximizes the generalized ratio, as shown in Table 3 across different values of N used for calculating the generalized ratio.

Table 3: Portfolio Allocation into the S&P500 using the CRRA Generalized Ratio

N	2	3	4	5	6	20
CRRA(1)	147.66%	145.65%	143.35%	143.15%	143.06%	143.03%
CRRA(2)	73.83%	73.07%	72.47%	72.44%	72.43%	72.43%
CRRA(3)	49.22%	48.77%	48.47%	48.46%	48.45%	48.45%
CRRA(4)	36.92%	36.6%	36.41%	36.4%	36.4%	36.4%
CRRA(5)	29.53%	29.29%	29.15%	29.15%	29.15%	29.15%

Explanation: CRRA(X) shows the optimal allocation into the S&P500, as a percentage of wealth, where X is the coefficient of risk aversion and N is the largest adjusted cumulant used in the shown calculation.

Notice that, in this example, the optimal allocation picked by the generalized ratio with $N = 20$ terms is not so different than the optimal allocation that would have been picked by the Sharpe ratio with just $N = 2$ terms. The reason is that S&P500 is almost Normally distributed at monthly frequency.

[16]To ensure that our results are not be driven by errors in the optimization engine, we use conservative (but slow) globally stable methods, including grid search and simplex. Furthermore, as discussed in the text, in several of our examples, by construction, we are able to verify our calculations using an alternative method.

However, one must still calculate the generalized ratio at larger values of N to verify that apparently small deviations from Normality in a given data set do not have a material impact on the optimal asset allocation. Despite Samuelson's (1970) conjecture about the approximation value of mean-variance analysis, there is no mapping between a distribution's goodness of fit to Normality and the sup norm error created by using the Sharpe ratio.

We can loosely "verify" the accuracy of our generalized ranking calculation by performing simulations on the original investor problem (1). Recall that the generalized ranking measure is calculated based on knowing only the translated moments of the underlying data. For the original investor problem (1), however, we need to know the actual risk distribution in order to integrate the expectation operator. Since, we don't have that information, we simply assume that the "true" distribution is given by the histogram of our data, which we then sample 100,000 times. Of course, this assumption could, in general, produce considerable error because it effectively eliminates the latent tails of the distribution, which could be especially problematic with non-Normal risk. In the case of broad S&P500 index, however, this effect appears to be small. Table 4 shows the results from this simulation analysis. Notice that the results are very close to calculations produced by the Generalized Ratio with $N = 20$.

Table 4: S&P500 – Simulation Results using the Investor Problem (1)

	Optimal Allocation
CRRA(1)	143.03%
CRRA(2)	72.41%
CRRA(3)	48.45%
CRRA(4)	36.4%
CRRA(5)	29.15%

Explanation: CRRA(X) shows the distribution-preference CRRA ranking, where X is the coefficient of risk aversion. CRRA results expressed as a percentage of wealth to be invested in the fund.

Table 5 shows the results for the approximate risk measure derived in Section 5.[17] Notice that the measure does fairly well at low values of risk aversion but performs poorly at higher values. Interesting, the approximation formula actually performs worse as more higher order terms are added (i.e., N grows). This example shows that approximation formulas run the risk of appearing to detect higher-order effects that don't actually exist, even in a sample that is already pretty close to being Normally distributed.

6.3 Fund-Level Reporting

Sharpe ratios are routinely reported for mutual and private funds in order to provide investors with guidance about how well the fund is performing relative to the risk being taken. A single Sharpe Ratio is typically provided for each fund (along with other information, of course).

As shown in Corollary 2, however, a single measure is not consistent with non-Normally distributed risk because all valid ranking measures must at least be conditioned on the preferences of investors. Moreover, the Sharpe ratio might incorrectly rank these funds. Accordingly, to produce useable information in an environment with non-Normal risk, investment managers need to produce a vector of generalized risk ratios conditional on preference parameters, using a sufficiently large value of N. As a

[17] At a value of $N = 2$, the results shown in Tables 3 and 5 should exactly match since both the generalized and approximate measures are equal to the Sharpe Ratio. The small differences come from the grid size used in the global optimizer.

Table 5: S&P500 – The Approximate Ratio

N	2	3	4	5	6	20
CRRA(1)	147.42%	143.64%	142.68%	142.47%	142.4%	142.39%
CRRA(2)	73.71%	68.03%	66.13%	65.59%	65.4%	65.34%
CRRA(3)	49.14%	41.57%	38.39%	37.31%	36.88%	36.71%
CRRA(4)	36.86%	27.39%	22.62%	20.74%	19.87%	19.47%
CRRA(5)	29.48%	18.13%	11.45%	8.43%	6.88%	6.07%

Explanation: CRRA(X) shows the optimal allocation into the S&P500, as a percentage of wealth, where X is the coefficient of risk aversion and N is the largest adjusted cumulant used in the shown calculation.

Table 6: Fund Ranking: The S&P500 Index as an Example Fund

$N = 20$	Investor Tolerance	Generalized Ranking Ratios
CRRA(1)	**"Growth"**	**0.00189**
CRRA(2)		0.00095
CRRA(3)	**"Moderate"**	**0.00063**
CRRA(4)		0.00048
CRRA(5)	**"Conservative"**	**0.00038**

Explanation: Hypothetical rankings for CRRA(X), where X is the coefficient of risk aversion with $N = 20$ adjusted cumulants.

practical matter, that is easy to do. Using the data described in the last subsection as an example, Table 6 shows the generalized risk ratios that could be reported for a hypothetical S&P500 indexed fund.[18] Ratios are shown for CRRA utility – a fairly standard investor utility function – across a range of risk tolerances, say 1 ("Growth"), 3 ("Moderate") and 5 ("Conservative"). The ratios for the risk tolerances of 2 and 4 are also shown for completeness.

Using this information, investors can now accurately compare funds based on their individual risk tolerance. Assessing an investor's risk tolerance in the presence of non-Normal risk is both necessary (by our impossibility theorem) and by no means far-fetched. Under the Uniform Securities Act, financial advisors must assess the risk tolerance of each client. Moreover, virtually all of the online broker-dealers offer "Investor Profile Questionnaires" that allows investors to determine their "risk tolerance."[19] Investor tolerances are usually broken down into 3 to 5 categories, similar to those shown in Table 6.

6.4 Multi-Asset Class Allocation

Now consider a multi-asset portfolio with underlying risk that is clearly non-Normally distributed. The following example is borrowed from the interesting papers by Goetzmann et al (2002) and Ingersoll et al (2007). The underlying stock follows a geometric Brownian motion with today's price normalized to

[18] As discussed in Section 3.4.2, the ranking measure is not necessarily monotone in the percent allocations invested into the fund (as shown in Table 4). Hence, the percent allocation is not a valid ranking measure. Instead, the generalized ratio itself should be shown.

[19] For example, Vanguard's Questionnaire can be found here: https://personal.vanguard.com/us/FundsInvQuestionnaire. [Last checked, August, 2013]. Schwab, Fidelity, and other major online broker-dealers offer similar surveys.

\$1 per share. The price in 1 year, therefore, is

$$S = \exp\left((\mu - \frac{1}{2}\sigma^2) + \sigma z\right),$$

where z is the standard Normal random variable. We set $\mu = 0.10$, $r_f = 0.05$, and $\sigma = 0.15$, which are reasonable values at an annual frequency, implying an annual equity premium of 5%.

In addition to holding a share of this stock, now include a European put option with a \$0.88 strike price and a European call option with \$1.12 strike price, both maturing in 1 year. Using the Black-Scholes-Merton model, the price for these put and call options today are \$0.0079 and \$0.0345, respectively. An investor can buy or sell either option in varying amounts. For example, a "12% collar" per share can be constructed by buying a put option, which protects against the stock price falling more than 12%, while selling a call option that gives away gains above 12%.[20]

Denote (a_1, a_2) as the allocation in put and call options, respectively. A positive value denotes buying the option while a negative value means selling (a writer). Our objective, therefore, is to determine the optimal buy/sell amounts of each option per share of stock.

The values of $a_1 = -1.3739$ and $a_2 = -0.5807$ maximize the Sharpe ratio, as shown in Table 7 for the case of $N = 2$. (Recall that the level of risk aversion is irrelevant for the Sharpe ratio, which is the reason that the values in Table 7 are identical across different risk preferences when $N = 2$.) In other words, the Sharpe ratio suggests a strong amount of put and call writing. In effect, the Sharpe ratio recommends that the investor's optimal position is short in volatility, thereby paying out for stock price swings in excess of 12% in either direction, but collecting the option premium as the reward.

Our generalized ratio, however, suggests that much less shorting of each option is optimal, as shown in Table 7 with $N = 20$. For example, at CRRA = 3 and $N = 20$, $a_1 = -0.160$ and $a_2 = -0.076$. The premium income collected from shorting the put equals \$0.0079(0.16) = \$0.00126 while the premium collected on shorting the call is \$0.0345(0.077) = \$0.00266, for a total of \$0.00392.[21]

In this example, we can also verify the generalized ratio calculations. By construction, we know the functional form for the risk distribution. Hence, we can alternatively use simulation analysis with many random draws, based on the original investor problem (1), to derive the optimal holdings of the put and call options. These results are also shown in Table 7 under the label of "Simulation." Notice that they closely match the results from the generalized ratio with $N = 20$.

Table 7: Optimal (Put, Call) Positions Per Share Held: The Generalized Ratio

N	2	3	5	20	Simulation
CARA	(-1.374, -0.581)	(-1.129,0.902)	(-0.402,-0.229)	(-0.390, -0.239)	(-0.390,-0.239)
CRRA(3)	(-1.374,-0.581)	NaN	(-0.482,0.529)	(-0.160,-0.077)	(-0.160,-0.076)
CRRA(4)	(-1.374,-0.581)	(-0.792,0.233)	(-0.317,0.065)	(-0.214,-0.120)	(-0.214,-0.120)
CRRA(5)	(-1.374,-0.581)	(-0.704,0.280)	(-0.310,-0.040)	(-0.247,-0.145)	(-0.247,-0.145)

Explanation: The optimal (put, call) allocation. For CRRA(X), X is the coefficient of risk aversion. N is the largest adjusted cumulant used in the shown calculation.

[20] The 12% value was arbitrarily chosen in this example to give the investor some upside above the mean, which is fairly common practice among those using collars. However, larger or smaller collar bounds could be used.

[21] Ingersoll et al (2007) develop a manager manipulation-proof measure around the (0,0) position. In contrast, we focus on the allocation that maximizes the investor expected utility problem, which will produce some demand for the new securities at fair prices since these securities are not spanned in discrete time.

7 Conclusions

The Sharpe ratio correctly ranks risky investments, consistent with the original investor problem, if risks are Normally distributed. Considerable past effort has been made to develop new measures that are robust to non-Normally distributed risks, which emerge with “fat tails,” modern trading strategies and the modern extensive use of financial derivatives in hedging portfolio risk. Some of this effort has started with the original investor problem and added some higher moments under fairly strong restrictions on the risk distribution and/or utility function. Even more ranking measures have been developed that satisfy certain mathematical properties or are purely ad hoc. Those measures, however, do not map back to an original investor problem, making their interpretation unclear. Not surprisingly, the Sharpe ratio, therefore, remains the gold standard in the industry, despite its lack of robustness to more general risk distributional assumptions.

This paper derives a generalized ranking measure that is valid under a broad admissible utility-probability space and yet preserves wealth separation for the broad HARA utility class. Our ranking measure can be used with non-Normal distributions. Because it can also pairwise compare composite risks following different distributions, it can also serve as the foundation for multi-asset class portfolio optimization, thereby replacing the mixture of other measures that are currently being used in industry. We demonstrate that the generalized ratio can produce very different optimal allocations than the Sharpe ratio, especially in the context of financial derivatives and other securities that produce non-Normal distributions. Along the way, we prove a key impossibility theorem: any ranking measure that is valid at non-Normal “higher moments” cannot generically be free from investor preferences. But, as a matter of practice, we demonstrate how the generalized ratio can be easily presented at the fund level for different risk tolerances, which already must be legally assessed by financial advisors for each investor.

References

[1] Agarwal, V. and Naik, N. Y. (2004). “Risk and Portfolio Decisions Involving Hedge Funds," Review of Financial Studies, 17 (1): 63-98.

[2] Artzner, P., Delbaen, F., Eber, J.M. and Heath, D. (1999). “Coherent Measures of Risk". Mathematical Finance 9 (3): 203-228.

[3] Aumann, R. J., and Roberto S. (2008). “An economic index of riskiness.” Journal of Political Economy, 116: 810-836.

[4] Barro, R. J. (2009). “Rare disasters, asset prices, and welfare costs.” American Economic Review, 99: 243-264.

[5] Black, F, and Scholes, M (1973). “The Pricing of Options and Corporate Liabilities.” Journal of Political Economy 81 (3): 637–654.

[6] Borch, K (1969). “A Note on Uncertainty and Indifference Curves,” Review of Economic Studies, 36: 1-4.

[7] Brandt, M.W., Goyal, A., Santa-Clara, P. and Stroud. J.R. (2005). “A simulation approach to dynamic portfolio choice with an application to learning about return predictability.” Review of Financial Studies, 18(3): 831.

[8] Brandt, M.W. (2010). “Chapter 5: Portfolio Choice Problems,” In Y. Ait-Sahalia and L.P. Hansen (eds.), Handbook of Financial Econometrics, Volume 1: Tools and Techniques, North Holland: 269-336.

[9] Brooks, C. and Kat, H. M. (2002). “The Statistical Properties of Hedge Fund Index Returns and Their Implications for Investors”, Journal of Alternative Investments, 5 (2): 26–44.

[10] Chamberlain, Gary. (1983). “A Characterization of the Distributions that Imply Mean-Variance Utility Functions.” Journal of Economic Theory, 29: 185 - 201.

[11] Cherny, A. (2003). “Generalized Sharpe Ratios and Asset Pricing in Incomplete Markets", European Finance Review ,7 (2): 191-233.

[12] Dávila, E. (2010). “Myopic Portfolio Choice with Higher Cumulants.” Working Paper, Harvard University.

[13] Dowd, K. (2000).“Adjusting for Risk: An Improved Sharpe Ratio", International Review of Economics and Finance, 9: 209-222.

[14] Durrett, R. (2005). “Probability: Theory and Examples 3rd Edition”, Thomson Learning^{TM}

[15] Fama, E. (1965). “The Behaviour of Stock-Market Prices", Journal of Business, 38: 34-105.

[16] Feldstein, M.S., (1969). “Mean-Variance Analysis in the Theory of Liquidity Preference and Portfolio Selection,” Review of Economic Studies, 36: 5-12.

[17] Foster, D.P., and Hart, S. (2009). “An Operational Measure of Riskiness.” Journal of Political Economy, 117: 785-814.

[18] Gabaix, Xavier (2012). “Variable Rare Disasters: An Exactly Solved Framework for Ten Puzzles in Macro-finance.” Quarterly Journal of Economics, 127(7): 645-700.

[19] Gourio, F (2012). “Disaster Risk and Business Cycles.” American Economic Review, 102(6): 2734-66.

[20] Goetzmann, W., Ingersoll, J., Spiegel, M., and Welch, I. (2002). “Sharpening Sharpe Ratios", NBER Working Paper #9116.

[21] Hart, S. (2011). “Comparing Risks by Acceptance and Rejection.” Journal of Political Economy, 119: 617-638.

[22] Hodges, S. (1998). “A Generalization of the Sharpe Ratio and its Applications to Valuation Bounds and Risk Measures,” Working Paper, Financial Options Research Centre, University of Warwick.

[23] Ingersoll, J.E. (1987). *Theory of financial decision making*. Rowman and Littlefield Studies in Financial Economics.

[24] Ingersoll, J., Spiegel, M., and Goetzmann, W. (2007). “Portfolio Performance Manipulation and Manipulation-proof Performance Measures", Review of Financial Studies, 20 (5): 1503-1546.

[25] Jensen, M.C. (1968). “The Performance of Mutual Funds in the Period 1945-1964,” Journal of Finance 23: 389-416.

[26] Jurczenko, E., and Maillet, B. (2006). “Theoretical Foundations of Asset Allocations and Pricing Models with Higher-order Moments", Jurczenko, E., and Maillet, B. (eds), *Multi-moment Asset Allocation and Pricing Models*, John Wiley & Sons, 233 p.

[27] Kane, M. (1982). “Skewness preference and portfolio choice.” Journal of Financial and Quantitative Analysis, 17(1): 15 - 25.

[28] Kaplan, P.D. and Knowles, J.A. (2004). “Kappa: A Generalized Downside Risk-Adjusted Performance Measure", Journal of Performance Measurement, 8 (3): 42-54.

[29] Kraus, A., and Litzenberger, R.H. (1976). “Skewness preference and the valuation of risk assets.” Journal of Finance 31(4): 1085–1100.

[30] Kroll, Y., Levy, H. and Markowitz, H.M. (1984). “Mean-Variance Versus Direct Utility Maximization.” The Journal of Finance 39(1): 47 - 61.

[31] Leland, H.E. (1999). “Beyond Mean-Variance: Risk and Performance Measurement in a Nonsymmetrical World", Financial Analysts Journal 1: 27 - 36.

[32] Levy, H. and Markowitz, H. (1979). “Approximating expected utility by a function of mean and variance.” American Economics Review 69: 308–317.

[33] Lukacs, E. (1970). Characteristic Functions (2nd Edition), Griffin, London.

[34] Malkiel, B.G. and Saha, A. (2005). “Hedge Funds: Risk and Return,” Financial Analysts Journal, Vol. 61(6): 80 - 88.

[35] Markowitz, H. (1952). “Portfolio Selection,” The Journal of Finance, 7: 77 - 91.

[36] Markowitz, H. (1959). *Portfolio Selection: Efficient Diversification of Investments*. John Wiley & Sons: New York.

[37] Merton, R. (1969) “Lifetime portfolio selection under uncertainty: The continuous-time case,” Review of Economics and Statistics 51: 247-257.

[38] Modigliani, F. (1997). “Risk-Adjusted Performance.” Journal of Portfolio Management (Winter): 45–54.

[39] Owen, J., and Rabinovitch, R. (1983). “On the class of elliptical distributions and their applications to the theory of portfolio choice.” Journal of Finance, 38(3): 745–752.

[40] Pierro, M. D. and Mosevich, J. (2011). “Effects of skewness and kurtosis on portfolio rankings,” Quantitative Finance, 11(10): 1449 - 1453.

[41] Roy, A.D. (1952). “Safety First and the Holding of Assets.” Econometrica 20: 431–450.

[42] Samuelson, P. (1970). “The Fundamental Approximation Theorem of Portfolio Analysis in Terms of Means, Variances and Higher Moments.” The Review of Economic Studies, 37(4): 537–542.

[43] Scott, R., and Horvath, P. (1980). “On the Direction of Preference for Moments of Higher Order than Variance.” Journal of Finance, 35: 915–919.

[44] Shadwick, W. F. and Keating, C. (2002). “A Universal Performance Measure", Journal of Performance Measurement, 6 (3): 59 - 84.

[45] Sharpe, W.F. (1964). “Capital asset prices: A theory of market equilibrium under conditions of risk.” Journal of Finance 19: 425 - 442.

[46] Sharpe, W.F. (1966). “Mutual Fund Performance,” Journal of Business 39: 119–138.

[47] Sortino, F.A. and Price, L.N. (1994). “Performance Measurement in a Downside Risk Framework,” Journal of Investing 3(3): 59 - 65.

[48] Spurgin, R.B. (2001). “How to Game Your Sharpe Ratio,” Journal of Alternative Investments 4(3): 38 - 46.

[49] Tobin, J. (1958). “Liquidity Preference as Behaviour Towards Risk,” Review of Economic Studies 25: 65 - 86.

[50] Tobin, J. (1969). “Comment on Borch and Feldstein,” Review of Econonmic Studies, 36: 13 - 14.

[51] Wachter, J. (2012). “Can Time-Varying Risk of Rare Disasters Explain Aggregate Stock Market Volatility?,” Journal of Finance 68(3): 987-1035.

[52] Zakamouline, V. and Koekebakker, S. (2009). “Portfolio Performance Evaluation with Generalized Sharpe Ratios: Beyond the Mean and Variance", Journal of Banking and Finance 33(7): 1242-1254.

Proofs

Theorem 1

Denote that $Z_\alpha = \frac{Y_\alpha - \mathbb{E}Y_\alpha}{\sqrt{Var(Y_\alpha)}}$, by our assumption Z_α doesn't depend on α, thus we can ignore the subscript: $Z = Z_\alpha$. The FOC of the maximization problem is

$$\mathbb{E}u'(w(1+r)+a^*(Y_\alpha - r))(Y_\alpha - r) = 0$$

in terms of Z:

$$\mathbb{E}u'(w(1+r)+a^*(Z\sqrt{Var(Y_\alpha)}+\mathbb{E}Y_\alpha - r))(Z\sqrt{Var(Y_\alpha)}+\mathbb{E}Y_\alpha - r) = 0$$

i.e.

$$\mathbb{E}u'\left(w(1+r)+a^*\sqrt{Var(Y_\alpha)}(Z+\frac{\mathbb{E}Y_\alpha - r}{\sqrt{Var(Y_\alpha)}})\right)(Z+\frac{\mathbb{E}Y_\alpha - r}{\sqrt{Var(Y_\alpha)}}) = 0$$

Given information about u, w, r, the solution $a^*\sqrt{Var(Y_\alpha)}$ only depends on $\frac{\mathbb{E}Y_\alpha - r}{\sqrt{Var(Y_\alpha)}}$ and Z. Since we assume Z_α doesn't depend on parameters α, we can write $a^*\sqrt{Var(Y_\alpha)}$ as a function of $\frac{\mathbb{E}Y_\alpha - r}{\sqrt{Var(Y_\alpha)}}$. Let us assume

$$a^*\sqrt{Var(Y_\alpha)} = g\left(\frac{\mathbb{E}Y_\alpha - r}{\sqrt{Var(Y_\alpha)}}\right)$$

Then

$$\begin{aligned}
&\max_a \mathbb{E}u(w(1+r)+a(Y_\alpha - r)) = \mathbb{E}u(w(1+r)+a^*(Y_\alpha - r)) \\
&= \mathbb{E}u(w(1+r)+a^*(Z\sqrt{Var(Y_\alpha)}+\mathbb{E}Y_\alpha - r)) \\
&= \mathbb{E}u\left(w(1+r)+a^*\sqrt{Var(Y_\alpha)}(Z+\frac{\mathbb{E}Y_\alpha - r}{\sqrt{Var(Y_\alpha)}})\right) \\
&= \mathbb{E}u\left(w(1+r)+g\left(\frac{\mathbb{E}Y_\alpha - r}{\sqrt{Var(Y_\alpha)}}\right)(Z+\frac{\mathbb{E}Y_\alpha - r}{\sqrt{Var(Y_\alpha)}})\right) \\
&= f_{(u,w,r)}\left(\frac{\mathbb{E}Y_\alpha - r}{\sqrt{Var(Y_\alpha)}}\right)
\end{aligned}$$

In addition, since Z is symmetric, if Y_1 and Y_2 produce opposite value of Sharpe ratio, i.e.

$$\frac{\mathbb{E}Y_1 - r}{\sqrt{Var(Y_1)}} = -\frac{\mathbb{E}Y_2 - r}{\sqrt{Var(Y_2)}}$$

The optimal allocation in these two cases would be opposite too, i.e.

$$a_1^*\sqrt{Var(Y_1)} = -a_2^*\sqrt{Var(Y_2)}$$

so $f_{(u,w,r)}$ is an even function. Let $\bar{f}(x) = f(x^2)$, clearly that $\bar{f}$ is an increasing function. Thus the term $\max_a \mathbb{E}u\left(w(1+r_f)+a(Y-r_f)\right)$ is an increasing function of $\left(\frac{\mathbb{E}Y - r_f}{\sqrt{Var(Y)}}\right)^2$.

Lemma 1

The uniqueness is because the maximization problem is concave, i.e. the second order derivative is

$$Eu''(w(1+r)+a(Y-r))(Y-r)^2 < 0$$

Let $h(a) = Eu'(w(1+r)+a(Y-r))(Y-r)$ be the first order derivative. It is easy to see that sign of $h(0)$ is same as sign of $\mathbb{E}Y - r$. So if $\mathbb{E}Y > r$,then $h(0) > 0$. We can conclude that the root of $h(a)$must be on right side of h because h is a decreasing function. Similarly we can show $h(0) = 0$ if $\mathbb{E}Y = r$ and $h(0) < 0$ if $\mathbb{E}Y < r$.

Lemma 2

On the complex plan, we can draw a small circle Γ around x_0 so that f have unique complex solution x_0 on Γ. Denote $\gamma = \partial\Gamma$ is the boundary of Γ. By Cauchy's Theorem, we have

$$\frac{1}{2\pi i}\oint_\gamma \frac{f'}{f}\,dz = 1$$

and

$$\frac{f_N'}{f_N} \to \frac{f'}{f}.$$

The fact that $\frac{1}{2\pi i}\oint_\gamma \frac{f_N'}{f_N}\,dz$ is always an integer, we conclude that for sufficient large N

$$\frac{1}{2\pi i}\oint_\gamma \frac{f_N'}{f_N}\,dz = 1$$

In other word, f_N has unique solution on Γ. In addition, since f_N is a polynomial with real coefficients, we can conclude that this root is a real number as real polynomials have conjugate complex roots.

Now we show that the unique solution of f_N on Γ is the smallest absolute root of f_N on the complex plane. We show by contradiction, i.e. for any N, there is $n > N$ so that f_n has a root that is smaller in absolute value. First, they are uniform bounded, say $2|x_0|$. Then, since it is a bounded convergent subsequence, we know that we can find a convergent subsequence and it must converge to x_0. This is an contradiction because those points are outside Γ, meaning they have a positive distance from x_0, resulting in impossibility of converging to x_0.

Example 6

Suppose utility is CARA and the underlying distribution is a discrete distribution, then we want to show that the Taylor expansion has infinite convergence radius, i.e. regularity condition always holds in this case. Suppose the underlying distribution is characterized by

$$\{x_1, p_1; x_2, p_2, ; \cdots ; x_m, p_m\}$$

and $A = \max\{|x_i - r_f|\} = |x_j - r_f|$. Then one can show that the t-moments $t_n \approx A^n$ for sufficient large n for the reason below. Then using familiar convergence radius formula, we obtain convergent radius

$$\limsup_k k/A = \infty$$

To calculate t_n, we have
$$p_j A^n \leq t_n = \sum_i p_i (x_i - r_f)^n \leq \sum_i p_i A^n = A^n.$$

Theorem 2

For HARA utility, $u(w) = \frac{\rho}{1-\rho}\left(\frac{\lambda w}{\rho} + \phi\right)^{1-\rho}$,

$$\begin{aligned}
&\max_a \mathbb{E}u\left(w(1+r_f) + a(Y - r_f)\right) \\
&= u(w_r) + \lambda\left(\frac{\lambda w_r}{\rho} + \phi\right)^{-\rho} \sum_{n=1}^{\infty} (\rho)\cdots(\rho+n-2)\frac{t_n^Y}{(n)!}\left(-\frac{\lambda}{\rho}\frac{a}{\frac{\lambda w_r}{\rho} + \phi}\right)^n \\
&= u(w_r) - \lambda\left(\frac{\lambda w_r}{\rho} + \phi\right)^{-\rho} \sum_{n=1}^{\infty} \frac{b_n t_n^Y}{(n)!} z^n
\end{aligned}$$

So $q_H(t_n^Y, b_n) = -\sum_{n=1}^N \frac{b_n t_n^Y}{n!} z_{N,Y}^n$ is the ranking function.

Theorem 3

We show by contradiction. Consider

$$Y = \begin{cases} k\% & \text{w.p. } p \\ -1\% & \text{w.p. } 1-p \end{cases}$$

Suppose investor's utility function is $u(w) = \frac{w^{1-\gamma}}{1-\gamma}$. Without loss of generality, assume initial wealth $w_0 = 1$, then investor solves following problem

$$\max_a \mathbb{E}u(1 + aY) = pu(1 + ak/100) + (1-p)u(1 - 0.01a)$$

FOC gives
$$pku'(1 + ak/100) = (1-p)u'(1 - 0.01a)$$

So we have
$$a^* = \frac{100(1 - (\frac{pk}{1-p})^{-1/\gamma})}{(1 + k * (\frac{pk}{1-p})^{-1/\gamma})}$$

The maximal value is then $\mathbb{E}u(1 + a^*Y) = pu(1 + a^*k) + (1-p)u(1 - 0.01a^*)$. Specifically, consider following example
$$Y_1 = \begin{cases} 1.6\% & \text{w.p. } 0.77 \\ -1\% & \text{w.p. } 0.23 \end{cases}$$

$$Y_2 = \begin{cases} 1.3\% & \text{w.p. } 0.81 \\ -1\% & \text{w.p. } 0.19 \end{cases}$$

Investors A and B are both CRRA with $\rho = 2$ and $\rho = 100$, respectively. Then

Investor	Y_1	Y_2	ranking
A	-0.8472	-0.8485	$Y_1 > Y_2$
B	-0.00722	-0.00717	$Y_1 < Y_2$

This implies there can not have distribution only ranking measure as if there is, we should expect investor A and B has same preference over Y_1 and Y_2.

Lemma 3

Since $e^{\psi(\theta,t)} = \mathbb{E}e^{\theta X_t}$, then the cumulant-generating function

$$g(\theta,t) = \log(\mathbb{E}e^{\theta X_t}) = \psi(\theta,t).$$

Denote k_n the n-th cumulant. Then

$$k_n = \frac{\partial g^n(\theta,1)}{\partial \theta^n}|_{\theta=0} = \frac{\partial \psi^n(\theta,1)}{\partial \theta^n}|_{\theta=0}.$$

So for $n \geq 2$,

$$k_n = \nu_n \sigma^n$$

Theorem 4

In general, for random variable Y, we denote μ_1 the mean and for $k \geq 2$,

$$\mu_k = \mathbb{E}(Y - \mathbb{E}Y)^k \text{ and } \xi_k = \frac{\mu_k}{\mu_2^{\frac{k}{2}}}$$

So ξ_3 represents the skewness and ξ_4 the kurtosis. We write S, K to represent skewness and kurtosis for simplicity. We also write $\mu_2 = \sigma^2$. We have

- $\nu_3 = \xi_3$
- $\nu_4 = \xi_4 - 3$
- $\nu_5 = \xi_5 - 5\xi_3$
- $\nu_6 = \xi_6 - \frac{15}{2}(\xi_4 - 3) - 10\xi_3^2 - 15 = \xi_6 - \frac{15}{2}\nu_4 - 10\nu_3^2 - 15$
- $\nu_7 = \xi_7 - \frac{21}{2}(\xi_5 - 5S) - \frac{35}{2}(K-3)S - \frac{175}{3}S = \xi_7 - \frac{21}{2}\nu_5 - \frac{35}{2}\nu_4\nu_3 - \frac{175}{3}\nu_3$

For integer n, there are numbers of ways to write it as sum of positive integers that greater than 1. For example, we can write

- 7=7
- 7=5+2
- 7=4+3
- 7=3+2+2

Those four ways to decomposing 7 matches the terms in ν_7 (noting that $\nu_2 = 1$). For a particular decomposition of n

$$n = i_1 + i_2 + \cdots + i_k$$

where $i_1 \geq i_2 \cdots \geq i_k \geq 2$. For $k \geq 2$, there is a corresponding term in ν_n that is $\nu_{i_1} \cdots \nu_{i_k}$ and the coefficient is

$$\frac{\binom{n}{i_1}\binom{n-i_1}{i_2}\cdots\binom{n-i_1-\cdots-i_{k-1}}{i_k}}{k!}$$

Thus, we conclude that

$$\nu_n = \xi_n - \sum_{\substack{n = i_1 + i_2 + \cdots + i_k \\ i_1 \geq i_2 \cdots \geq i_k \geq 2}} \frac{\binom{n}{i_1}\binom{n-i_1}{i_2}\cdots\binom{n-i_1-\cdots-i_{k-1}}{i_k}}{k!} \nu_{i_1} \cdots \nu_{i_k}$$

Indeed let $Y = X_1 + X_2 + \cdots + X_m$, where X_i are i.i.d. Let's denote X for simplicity. Expand

$$\begin{aligned}
&\mathbb{E}(Y - \mathbb{E}Y)^n \\
=&\mathbb{E}(X_1 + X_2 + \cdots + X_m - \mathbb{E}(X_1 + X_2 + \cdots + X_m))^n \\
=&\mathbb{E}((X_1 - \mathbb{E}X_1) + (X_2 - \mathbb{E}X_2) + \cdots + (X_m - \mathbb{E}X_m))^n \\
=&\sum_{i_1+i_2+\cdots+i_m=n} \mathbb{E}(X_1 - \mathbb{E}X_1)^{i_1} \cdots (X_m - \mathbb{E}X_m)^{i_m} \\
=&m\mathbb{E}(X - \mathbb{E}X)^n + \sum_{\substack{i_1 + i_2 + \cdots + i_m = n \\ 2 \leq i_k < n,\ or\ i_k = 0}} \mathbb{E}(X_1 - \mathbb{E}X_1)^{i_1} \cdots (X_m - \mathbb{E}X_m)^{i_m}
\end{aligned}$$

Theorem 5

Suppose X_t is Levy Process, $Y = X_1$ then Y is infinitely divisible distribution with $X = X_{\frac{1}{n}}$. By Levy-Khinchine representation, we have

$$\mathbb{E}e^{i\theta X_t} = exp\left(bit\theta - \frac{1}{2}\sigma_0^2 t\theta^2 + t\int_{\mathbb{R}\setminus\{0\}} (e^{i\theta x} - 1 - i\theta x I_{|x|<1}) W(dx)\right)$$

where $b \in \mathbb{R}$, and I is the indicator function. The Levy measure W must be such that

$$\int_{\mathbb{R}\setminus\{0\}} min\{x^2, 1\} W(dx) < \infty$$

Denote

$$\phi(\theta, t) = bit\theta - \frac{1}{2}\sigma_0^2 t\theta^2 + t\int_{\mathbb{R}\setminus\{0\}} (e^{i\theta x} - 1 - i\theta x I_{|x|<1}) W(dx)$$

and

$$\psi(\theta, t) = bt\theta + \frac{1}{2}\sigma_0^2 t\theta^2 + t\int_{\mathbb{R}\setminus\{0\}} (e^{\theta x} - 1 - \theta x I_{|x|<1}) W(dx)$$

i.e $\phi(\theta, t) = \psi(i\theta, t)$ and $e^{\psi(\theta,t)}$ is the moments generating function of X_t. Suppose $e^{\psi(\theta,t)}$ are the moments generating function of Levy Process X_t, let $Y = X_1$, and σ is the standard deviation of Y, then we have

$$\nu_k(Y) = \frac{\frac{\partial^k \psi(\theta,1)}{\partial \theta^k}|_{\theta=0}}{\sigma^k}, \forall k \geq 2.$$

Theorem 6

The derivations in the text before the theorem was stated serves as proof.

Lemma 4

1. Suppose investor's preference is characterized by CARA and X_t is Levy Process. Let$Y = X_1$ then Y is infinitely divisible distribution with $X = X_{\frac{1}{n}}$. By Levy-Khinchine representation, we have

$$\mathbb{E}e^{i\theta X_t} = exp\left(bit\theta - \frac{1}{2}\sigma_0^2 t\theta^2 + t\int_{\mathbb{R}\setminus\{0\}}(e^{i\theta x} - 1 - i\theta x I_{|x|<1})W(dx)\right)$$

where $b \in \mathbb{R}$, and I is the indicator function. The Levy measure W must be such that

$$\int_{\mathbb{R}\setminus\{0\}} min\{x^2, 1\}W(dx) < \infty$$

Denote

$$\phi(\theta,t) = bit\theta - \frac{1}{2}\sigma_0^2 t\theta^2 + t\int_{\mathbb{R}\setminus\{0\}}(e^{i\theta x} - 1 - i\theta x I_{|x|<1})W(dx)$$

and

$$\psi(\theta,t) = bt\theta + \frac{1}{2}\sigma_0^2 t\theta^2 + t\int_{\mathbb{R}\setminus\{0\}}(e^{\theta x} - 1 - \theta x I_{|x|<1})W(dx)$$

i.e $\phi(\theta,t) = \psi(i\theta,t)$ and $e^{\psi(\theta,t)}$ is the moments generating function of X_t. From Theorem 5 we have,

$$\nu_k = \lim_{n\to\infty}\mathbb{E}\varepsilon^k n^{-k/2+1} = \frac{\mathbb{E}X_{\frac{1}{n}}^k}{\sigma^k}n = \frac{\frac{\partial^k\psi(\theta,1)}{\partial\theta^k}\,|_{\theta=0}}{\sigma^k}$$

Now suppose investor has CARA utility and as we mention before the risk aversion doesn't matter because of $b_n = 1$, so for simplicity, we assume $u(w) = -e^{-w}$.

$$\begin{aligned}
a^* &= arg\max_a \mathbb{E} - e^{-(w_r + a(Y-r))} \\
&= arg\max_a \mathbb{E} - e^{-w_r + ar - aY} \\
&= arg\max_a \mathbb{E} - e^{-w_r + ar}e^{-aY} \\
&= arg\max_a -e^{-w_r+ar}\mathbb{E}e^{-aY} \\
&= arg\max_a -e^{-w_r+ar}e^{\psi(-a,1)} \\
&= arg\min_a e^{ar+\psi(-a,1)} \\
&= arg\min_a ar + \psi(-a,1)
\end{aligned}$$

i.e. a^* is the unique solution of

$$r = \psi'(-a,1)$$

For adjusted cumulants performance measurement, we want to solve

$$\sum_{n=1}^{N} u^{(n)}(w_r)\frac{\nu_n}{(n-1)!}(\sigma a)^{n-1} = 0$$

Plug in $u(w) = -e^{-w}$ and $\nu_k = \frac{\frac{\partial^k \psi(\theta,1)}{\partial \theta^k}|_{\theta=0}}{\sigma^k}$ for $k \geq 2$ we get

$$\frac{\mu - r}{\sigma} + \sum_{n=2}^{N} \frac{\frac{\frac{\partial^n \psi(\theta,1)}{\partial \theta^n}|_{\theta=0}}{\sigma^n}}{(n-1)!}(-\sigma a)^{n-1} = 0$$

i.e.

$$\mu - r + \sum_{n=1}^{N-1} \frac{\frac{\partial^{n+1}\psi(\theta,1)}{\partial \theta^{n+1}}|_{\theta=0}}{n!}(-a)^n = 0$$

Notice that $\mu = \frac{\partial \psi(\theta,1)}{\partial \theta}|_{\theta=0}$, i.e.

$$r = \sum_{n=0}^{N-1} \frac{\frac{\partial^{n+1}\psi(\theta,1)}{\partial \theta^{n+1}}|_{\theta=0}}{n!}(-a)^n$$

By Taylor expansion we knows that the right hand side approaches $\psi'(-a,1)$. From Lemma 2, we have $a_N^* \to a^*$. In other word, the Adjusted Cumulant Ranking Measure is consistent with the maximized expected utility, therefore CARA is scalable.

1. For quadractic risk, since Sharpe ratio sqaured is a valid ranking measure and notice that

$$\frac{\mu \Delta t - r\Delta t}{\sigma \Delta t} = \frac{\mu - r}{\sigma}$$

2. For HARA and Possion, since it is one parameter distribution (λ), we can show that the maximized expected utility is an increasing function of $\lambda$when $\lambda \geq r$ and decreasing function when $\lambda < r$. When it is scaled to Δt, we have the maximized expected utility to be an increasing function $\lambda$when $\lambda \Delta t \geq r\Delta t$ and decreasing function when $\lambda \Delta t < r\Delta t$. So the maximized expected utility in the regular time and Δt time have the same ranking among all Possion distributions.

3. For Normal risk, the result is also from the fact that $\frac{\mu \Delta t - r\Delta t}{\sigma \Delta t} = \frac{\mu - r}{\sigma}$.

Theorem 7

First we claim that the adjusted cumulant ranking measure produces same ranking over time t and time Δt. In other words, $(Y, r) \geq_{ACRM} (Y', r) \Leftrightarrow (X, r\Delta t) \geq_{ACRM} (X', r\Delta t)$.

By Levy-Khinchine representation, we have in general

$$\nu_n^X = \nu_n^Y k^{\frac{n-2}{2}}$$

Now look at the equations that solve a^*. a^{*Y} solves

$$\sum_{n=1}^{N} u^{(n)}(w) \frac{\nu_n^Y}{(n-1)!}(\sigma a)^{n-1} = 0 \tag{15}$$

while a^{*X} solves

$$\sum_{n=1}^{N} u^{(n)}(w) \frac{\nu_n^X}{(n-1)!}(\sigma a)^{n-1} = 0 \tag{16}$$

Since $\nu_n^X = \nu_n^Y k^{\frac{n-2}{2}}$, so (16) is equivalent to

$$\sum_{n=1}^{N} u^{(n)}(w)\frac{\nu_n^Y k^{\frac{n-2}{2}}}{(n-1)!}(\sigma a)^{n-1} = 0$$

i.e.

$$k^{-\frac{1}{2}}\sum_{n=1}^{N} u^{(n)}(w)\frac{\nu_n^Y}{(n-1)!}(\sigma a\sqrt{k})^{n-1} = 0$$

Compared it to (15), we have

$$a^{*Y} = a^{*X}\sqrt{k}$$

Therefore, the left hand side of theorem

$$\begin{aligned}
&\sum_{n=1}^{N}\frac{u^{(n)}(w)}{u'(w)}\frac{\nu_n^{X_{k,1}}}{n!}(\sigma a^{*X_{k,1}})^n \\
&= \sum_{n=1}^{N}\frac{u^{(n)}(w)}{u'(w)}\frac{\nu_n^Y k^{\frac{n-2}{2}}}{n!}(\sigma a^{*Y}k^{-\frac{1}{2}})^n \\
&= \frac{1}{k}\sum_{n=1}^{N}\frac{u^{(n)}(w)}{u'(w)}\frac{\nu_n^Y}{n!}(\sigma a^{*Y})^n
\end{aligned}$$

Or

$$ACRM_N(X_{k,1},\frac{r_f}{k}) = \frac{1}{k}PM(Y,r_f)$$

where

$$ACRM_N(Y,r_f) = \sum_{n=1}^{N}\frac{u^{(n)}(w)}{u'(w)}\frac{\nu_n^Y}{n!}(\sigma a^{*Y})^n$$

Therefore, suppose we have infinitely divisible distributions Y and Y' in χ_m^N, we must have the following equivalence:

$$ACRM_N(Y,r_f) \geq ACRM_N(Y',r_f) \Leftrightarrow ACRM_N(X_{k,1},\frac{r_f}{k}) \geq ACRM_N(X'_{k,1},\frac{r_f}{k}), \forall k$$

Now above argument is also correct when $N = +\infty$, so we obtain

$$ACRM(Y,r_f) \geq ACRM(Y',r_f) \Leftrightarrow ACRM(X_{k,1},\frac{r_f}{k}) \geq ACRM(X'_{k,1},\frac{r_f}{k}), \forall k$$

Since u is scalable with respect to χ, we have $(Y,r) \geq_u (Y',r) \Leftrightarrow (X,r\Delta t) \geq_u (X',r\Delta t)$. By above derivation, we have $(X,r\Delta t) \geq_u (X',r\Delta t) \Leftrightarrow (X,r\Delta t) \geq_{ACRM} (X',r\Delta t)$, where $\geq_{ACRM}$ means adjusted cumulant ranking measure score function. From above claim, we have $(X,r\Delta t) \geq_{ACRM} (X',r\Delta t) \Leftrightarrow (Y,r) \geq_{ACRM} (Y',r)$. So we establish that $(Y,r) \geq_{ACRM} (Y',r) \Leftrightarrow (Y,r) \geq_u (Y',r)$.

Theorem 8

Suppose $\nu_n = c_{n,} \forall n \geq 3$. We need to solve

$$\sum_{n=1}^{2} u^{(n)}(w)\frac{\nu_n}{(n-1)!}(\sigma a)^{n-1} + \sum_{n=4,even}^{\infty} u^{(n)}(w)\frac{c_n}{(n-1)!}(\sigma a)^{n-1} = 0$$

Suppose we have$\sigma a = g(\nu_1)$. Then the let PM be the corresponding ranking furnction

$$PM = \sum_{n=1}^{2} u^{(n)}(w)\frac{\nu_n}{n!}(g(\nu_1))^n + \sum_{n=4,even}^{\infty} u^{(n)}(w)\frac{c_n}{(n-1)!}(g(\nu_1))^n$$

Then we have

$$\frac{\partial PM}{\partial \nu_1} = u'(w)g(\nu_1) + g'(\nu_1)\sum_{n=1}^{\infty} u^{(n)}(w)\frac{\nu_n}{(n-1)!}(\sigma a)^{n-1} = u'(w)g(\nu_1)$$

It it positive whenever ν_1 is poisitive and negative when ν_1 is negative. In addition, the measure is symmetric because $PM(\nu_1) = PM(-\nu_1)$. So Sharpe ratio is valid.

Lemma 5

This Lemma is similar to a result that is well known in the literature. See, in particular, Lukacs (1970).

Theorem 9

We use chain rule:

$$\begin{aligned}
&\frac{\partial \sum_{n=1}^{N} \frac{u^{(n)}(w_r)}{u'(w_r)}\frac{\nu_n}{n!}(\sigma a^*)^n}{\partial \nu_i} \\
=&\frac{\partial \sum_{n\neq i} \frac{u^{(n)}(w_r)}{u'(w_r)}\frac{\nu_n}{n!}(\sigma a^*)^n}{\partial \nu_i} + \frac{\partial \frac{u^{(i)}(w_r)}{u'(w_r)}\frac{\nu_i}{i!}(\sigma a^*)^i}{\partial \nu_i} \\
=&\sum_{n\neq i} \frac{u^{(n)}(w_r)}{u'(w_r)}\frac{\nu_n}{n!}\frac{\partial(\sigma a^*)^n}{\partial \nu_i} + \frac{u^{(i)}(w_r)}{u'(w_r)}\frac{1}{i!}(\sigma a^*)^i + \frac{u^{(i)}(w_r)}{u'(w_r)}\frac{\nu_i}{i!}\frac{\partial(\sigma a^*)^i}{\partial \nu_i} \\
=&\sum_{n=1}^{N} \frac{u^{(n)}(w_r)}{u'(w_r)}\frac{\nu_n}{n!}\frac{\partial(\sigma a^*)^n}{\partial \nu_i} + \frac{u^{(i)}(w_r)}{u'(w_r)}\frac{1}{i!}(\sigma a^*)^i \\
=&\sum_{n=1}^{N} \frac{u^{(n)}(w_r)}{u'(w_r)}\frac{\nu_n}{n!}n(\sigma a^*)^{n-1}\frac{\partial(\sigma a^*)}{\partial \nu_i} + \frac{u^{(i)}(w_r)}{u'(w_r)}\frac{1}{i!}(\sigma a^*)^i \\
=&\left(\sum_{n=1}^{N} \frac{u^{(n)}(w_r)}{u'(w_r)}\frac{\nu_n}{(n-1)!}(\sigma a^*)^{n-1}\right)\frac{\partial(\sigma a^*)}{\partial \nu_i} + \frac{u^{(i)}(w_r)}{u'(w_r)}\frac{1}{i!}(\sigma a^*)^i \\
=&\frac{u^{(i)}(w_r)}{u'(w_r)}\frac{1}{i!}(\sigma a^*)^i
\end{aligned}$$

Use lemma 2.1 and the fact that u has positive odd derivatives and negative even derivatives, we can easily get the conclusion:

- If $\mu > r$, it is increasing with respect to odd adjusted cumulants $\nu_3, \nu_5, \cdots$ and decreasing with respect to even adjusted cumulants $\nu_4, \nu_6, \cdots$.
- If $\mu < r$, it is decreasing with respect to odd adjusted cumulant $\nu_3, \nu_5, \cdots$ and increasing with respect to even adjusted cumulant $\nu_4, \nu_6, \cdots$.

Theorem 10

We have

$$\sigma a^* \approx -\frac{u'(w_r)}{u''(w_r)}\frac{\mu - r}{\sigma} + \sum_{n=3}^{N} c_n \frac{u^{(n)}(w_r)}{u'(w_r)}\nu_n$$

Plug into the first order condition and only keeps terms of ν_n and ignore higher terms:

$$\frac{u''(w_r)}{u'(w_r)}\sum_{n=3}^{N} c_n \frac{u^{(n)}(w_r)}{u'(w_r)}\nu_n + \sum_{n=3}^{N}\frac{u^{(n)}(w_r)}{u'(w_r)}\frac{\nu_n}{(n-1)!}\left(-\frac{u'(w_r)}{u''(w_r)}\frac{\mu - r}{\sigma}\right)^{n-1} = 0.$$

Since $p_n = \frac{u^{(n)}(w_r)}{u'(w_r)}$, then

$$p_2 \sum_{n=3}^{N} c_n p_n \nu_n + \sum_{n=3}^{N} p_n \frac{\nu_n}{(n-1)!}\left(\left(-\frac{1}{p_2}\frac{\mu - r}{\sigma}\right)^{n-1}\right) = 0.$$

This implies that

$$c_n = \frac{(-\frac{1}{p2})^n (\frac{\mu - r}{\sigma})^{n-1}}{(n-1)!}$$

So

$$\sigma a^* = -\frac{1}{p_2}\frac{\mu - r}{\sigma} + \sum_{n=3}^{N}\frac{(-\frac{1}{p2})^n (\frac{\mu - r}{\sigma})^{n-1}}{(n-1)!} p_n \nu_n.$$

Now we plug in σa^* to the expected utility

$$\begin{aligned}
&\max_a \mathbb{E} u(w_r + a(X - r)) \\
&= u(w_r) + \left(\sum_{n=1}^{N} u^{(n)}(w_r)\frac{\nu_n}{n!}(\sigma a^*)^n\right)\Delta t + o(\Delta t) \\
&= u(w_r) + u'(w_r)\left(\sum_{n=1}^{N} p_n \frac{\nu_n}{n!}\left(-\frac{1}{p_2}\frac{\mu - r}{\sigma} + \sum_{k=3}^{N}\frac{(-\frac{1}{p2})^k (\frac{\mu - r}{\sigma})^{k-1}}{(k-1)!} p_k \nu_k\right)^n\right)\Delta t + o(\Delta t) \\
&\approx u\left(w_r + \sum_{n=1}^{N} p_n \frac{\nu_n}{n!}\left(-\frac{1}{p_2}\frac{\mu - r}{\sigma} + \sum_{k=3}^{N}\frac{(-\frac{1}{p2})^k (\frac{\mu - r}{\sigma})^{k-1}}{(k-1)!} p_k \nu_k\right)^n \Delta t\right)
\end{aligned}$$

then we have the approximation performance measure in Nth moments given by

$$-\frac{(\frac{\mu - r}{\sigma})^2}{2p_2}\Delta t + \sum_{k=3}^{N}\frac{p_k \nu_k}{k!}\left(-\frac{\mu - r}{p_2 \sigma}\right)^k \Delta t$$

Rewrite in terms of adjusted cumulants:

$$-\frac{\nu_1^2}{2p_2} + \sum_{k=3}^{N}\frac{p_k \nu_k}{k!}\left(-\frac{\nu_1}{p_2}\right)^k$$

For Remark 6, notice that $b_n = \frac{p_n}{p_2^{n-1}}$ distribute the Δt, then we have

$$
\begin{aligned}
& -\frac{(\frac{\mu-r}{\sigma})^2}{2p_2}\Delta t + \sum_{k=3}^{N}\frac{p_k \nu_k}{k!}(-\frac{\mu-r}{p_2\sigma})^k \Delta t \\
&= -\frac{(\frac{\mu-r}{\sigma}\sqrt{\Delta t})^2}{2p_2} + \sum_{k=3}^{N}\frac{(-1)^k b_k \nu_k}{k! p_2}(\frac{\mu-r}{\sigma}\sqrt{\Delta t})^k (\Delta t)^{1-k/2} \\
&= -\frac{(\frac{\mu-r}{\sigma}\sqrt{\Delta t})^2}{2p_2} + \frac{1}{p_2}\sum_{k=3}^{N}\frac{(-1)^k b_k}{k!} SR^k (\nu_k (\Delta t)^{1-k/2}) \\
&= -\frac{SR^2}{2p_2} + \frac{1}{p_2}\sum_{k=3}^{N}\frac{(-1)^k b_k}{k!} SR^k (\nu_k (\Delta t)^{1-k/2}) \\
&= \frac{-1}{p_2}\left(\frac{SR^2}{2} + \sum_{k=3}^{N}\frac{(-1)^{k-1} b_k}{k!} SR^k (\nu_k (\Delta t)^{1-k/2})\right) \\
&\approx \frac{-1}{p_2}\left(\frac{SR^2}{2} + \sum_{k=3}^{N}\frac{(-1)^{k-1} b_k}{k!} SR^k \mathbb{E}\varepsilon^k\right)
\end{aligned}
$$