*Research Article*

# Inference for the Sharpe Ratio Using a Likelihood-Based Approach

## Ying Liu,[1] Marie Rekkas,[2] and Augustine Wong[3]

[1] *Department of Economics, York University, 4700 Keele Street, Toronto, ON, Canada M3J 1P3*

[2] *Department of Economics, Simon Fraser University, 8888 University Drive, Burnaby, BC, Canada V5A 1S6*

[3] *Department of Mathematics and Statistics, York University, 4700 Keele Street, Toronto, ON, Canada M3J 1P3*

Correspondence should be addressed to Marie Rekkas, mrekkas@sfu.ca

The Sharpe ratio is the prominent risk-adjusted performance measure used by practitioners. Statistical testing of this ratio using its asymptotic distribution has lagged behind its use. In this paper, highly accurate likelihood analysis is applied for inference on the Sharpe ratio. Both the one- and two-sample problems are considered. The methodology has $O(n^{-3/2})$ distributional accuracy and can be implemented using any parametric return distribution structure. Simulations are provided to demonstrate the method's superior accuracy over existing methods used for testing in the literature.

## 1. Introduction

The measurement of fund performance is an integral part of investment analysis. Investments are often ranked and evaluated on the basis of their risk-adjusted returns. Several risk-adjusted performance measures are available to money managers of which the Sharpe ratio is the most popular. Introduced by William Sharpe in 1966 [1], this ratio provides a measure of a fund's excess returns relative to its volatility. Expressed in its usual form, the Sharpe ratio for an asset with an expected return given by $\mu$ and standard deviation given by $\sigma$ is given by the following:

$$\text{SR} = \frac{\mu - R_f}{\sigma}, \tag{1.1}$$

where $R_f$ is the risk-free rate of return. From this expression, it is clear to see how this ratio provides a measure of a fund's excess return per unit of risk.

The Sharpe ratio has been extensively studied in the literature. The main criticism leveled against this measure concerns its reliance on only the first two moments of the returns distribution. If investment returns are normally distributed then the Sharpe ratio can be justified. On the other hand, if returns are asymmetric then it can be argued that the measure may not accurately describe the fund's performance as moments reflecting skewness and kurtosis are not captured by the ratio. To address this issue, several measures exist in the literature which integrate higher moments into the performance measure. The Omega measure is one such measure that uses all the available information in the returns distribution. Keating and Shadwick [2] provide an introduction to this measure. While various methods are available, they are also more complex and often very difficult to implement in practice. To gauge the trade-off between the attractiveness of such measures and their cost, Eling and Schuhmacher [3] compared the Sharpe ratio with 12 other approaches to performance measurement. Eling and Schuhmacher [3] focussed on the returns of 2,763 hedge funds. Hedge funds are known to have return distributions which differ significantly from the normal distribution and, as such, provide a rich and relevant environment for such a comparison. Their study concluded that the Sharpe ratio produced rankings that were largely identical to those obtained from the 12 other performance measures. In other words, the choice of performance measure did not matter much in terms of ranking funds. From a practical perspective, this imparts credibility to the use of the Sharpe ratio for ranking funds with highly nonnormal returns.

Our focus in this paper is on the statistical properties of the Sharpe ratio. While this ratio may be the most widely known and used measure of risk-adjusted performance for an investment fund, fund managers rarely (if ever) indicate any measure of associated statistical significance of their rankings. This point was made in Opdyke [4]. We use a likelihood-based statistical method to obtain highly accurate $p$ values and confidence intervals for evaluating the significance of the Sharpe ratio. While the methodology we use is applicable under *any* parametric distributional assumption, for expositional clarity we demonstrate the use of the method under the assumption of normal returns. We investigate inference for both the one-sample problem as well as the two-sample problem. We compare our method to currently existing methods in the literature. Simulations indicate the superior performance of the method compared to existing methods in the literature.

The paper is organized as follows. Section 2 provides an overview of the asymptotic distribution of the Sharpe ratio under various assumptions. Section 3 lays out the framework for the likelihood-based approach we will use for testing. Section 4 applies this methodology for inference on the Sharpe ratio. Examples and simulations for this one-sample case are then presented in Section 5. The two-sample problem is given in Section 6, with example and simulations presented in Section 7. Section 8 concludes.

## 2. Distribution of the Sharpe Ratio

Jobson and Korkie [5] and Lo [6] derived the asymptotic distribution of the Sharpe ratio assuming identically and independently normally distributed returns. Consider a fund with a return at time $t$ given by $R_t$, $t = 1, \ldots, T$. Further assume that the returns are identically and

independently distributed (IID) as $N(\mu, \sigma^2)$. If we let the vector $\theta = (\mu, \sigma^2)'$, then as given in Lo [6] we have the following asymptotically normal distribution for $\hat{\theta}$:

$$\sqrt{T}\left(\hat{\theta} - \theta\right) \longrightarrow N(0, V), \tag{2.1}$$

where

$$V \equiv \begin{pmatrix} \sigma^2 & 0 \\ 0 & 2\sigma^4 \end{pmatrix}, \tag{2.2}$$

and $\hat{\theta} = (\hat{\mu}, \hat{\sigma}^2)'$ is the maximum likelihood estimator (MLE) with $\hat{\mu}$ and $\hat{\sigma}^2$ given by the following formulas

$$\hat{\mu} = \frac{1}{T}\sum_{t=1}^{T} R_t,$$

$$\hat{\sigma}^2 = \frac{1}{T}\sum_{t=1}^{T}\left(R_t - \hat{\mu}\right)^2. \tag{2.3}$$

Lo's result is straightforward to prove. For large samples, by the central limit theorem, we have

$$\left(\hat{\theta} - \theta\right)'\left[\text{var}\left(\hat{\theta}\right)\right]^{-1}\left(\hat{\theta} - \theta\right) \longrightarrow \chi^2_k, \tag{2.4}$$

where $k = \dim(\theta)$ and

$$\text{var}\left(\hat{\theta}\right) = \begin{pmatrix} \text{var}(\hat{\mu}) & \text{cov}(\hat{\mu}, \hat{\sigma}^2) \\ \text{cov}(\hat{\mu}, \hat{\sigma}^2) & \text{var}(\hat{\sigma}^2) \end{pmatrix}. \tag{2.5}$$

See for instance, Cox and Hinkley [7] for this result. Given $\mu = (1,0)\theta$ and $\sigma^2 = (0,1)\theta$, we have

$$\frac{(\hat{\mu} - \mu)}{\sqrt{\text{var}(\hat{\mu})}} \longrightarrow N(0,1), \qquad \frac{(\hat{\sigma}^2 - \sigma^2)}{\sqrt{\text{var}(\hat{\sigma}^2)}} \longrightarrow N(0,1). \tag{2.6}$$

Given the Sharpe ratio is a function of $\theta$, the Delta method can be used to derive its asymptotic distribution. Let $\text{SR} = g(\theta)$ be a continuous function with Jacobian denoted by $\partial g(\theta)/\partial\theta$ and let the Sharpe ratio estimate be given by $\widehat{\text{SR}} = g(\hat{\theta})$. Applying a first-order Taylor expansion on $g(\hat{\theta})$ around the true parameter value $\theta$, we have

$$g\left(\hat{\theta}\right) \doteq g(\theta) + \left[\frac{\partial g(\theta)}{\partial\theta}\right]'\left(\hat{\theta} - \theta\right). \tag{2.7}$$

Rearranging terms and multiplying by $\sqrt{T}$ yields

$$\sqrt{T}\left(g\left(\hat{\theta}\right) - g(\theta)\right) \doteq \left[\frac{\partial g(\theta)}{\partial \theta}\right]' \sqrt{T}\left(\hat{\theta} - \theta\right). \tag{2.8}$$

Using the above result and applying the Delta method, we have the asymptotic distribution of $g(\hat{\theta})$:

$$\sqrt{T}\left(g\left(\hat{\theta}\right) - g(\theta)\right) \doteq \left[\frac{\partial g(\theta)}{\partial \theta}\right]' \sqrt{T}\left(\hat{\theta} - \theta\right) \longrightarrow N\left(0, \left[\frac{\partial g(\theta)}{\partial \theta}\right]' V \frac{\partial g(\theta)}{\partial \theta}\right). \tag{2.9}$$

Note that the Jacobian term is readily computed as follows:

$$\frac{\partial g(\theta)}{\partial \theta} = \begin{pmatrix} \dfrac{\partial g(\theta)}{\partial \mu} \\[2mm] \dfrac{\partial g(\theta)}{\partial \sigma^2} \end{pmatrix} = \begin{pmatrix} \dfrac{1}{\sigma} \\[2mm] -\dfrac{(\mu - R_f)}{2\sigma^3} \end{pmatrix}. \tag{2.10}$$

Combining terms yields the variance of the asymptotic distribution:

$$\left[\frac{\partial g(\theta)}{\partial \theta}\right]' V \frac{\partial g(\theta)}{\partial \theta} = 1 + \frac{(\mu - R_f)^2}{2\sigma^2} = 1 + \frac{1}{2}\text{SR}^2. \tag{2.11}$$

By the central limit theorem, the distribution of the Sharpe ratio assuming identically and independently distributed normal returns is therefore

$$\frac{\sqrt{T}\left(\widehat{\text{SR}} - \text{SR}\right)}{\sqrt{1 + (1/2)\widehat{\text{SR}}^2}} \longrightarrow N(0, 1). \tag{2.12}$$

From this distribution, $100(1-\alpha)\%$ confidence intervals for the Sharpe ratio can be constructed in the usual fashion:

$$\widehat{\text{SR}} \pm z_{\alpha/2}\sqrt{\frac{\left(1 + (1/2)\widehat{\text{SR}}^2\right)}{T}}, \tag{2.13}$$

where $z_{\alpha/2}$ is the $100 \times (1 - \alpha/2)$th percentile of the standard normal distribution.

The above derivation only holds under the IID normal assumption. Under the IID assumption but without assuming normally distributed returns, Mertens [8] derived the following asymptotic distribution of the Sharpe ratio:

$$\frac{\sqrt{T}\left(\widehat{\text{SR}} - \text{SR}\right)}{\sqrt{1 + (1/2)\widehat{\text{SR}}^2 - \widehat{\text{SR}} \cdot \hat{\gamma}_3 + \widehat{\text{SR}}^2 \cdot ((\hat{\gamma}_4 - 3)/4)}} \longrightarrow N(0, 1), \tag{2.14}$$

where $\gamma_3$ and $\gamma_4$ are defined as follows:

$$\gamma_3 = \frac{E(R_t - \mu)^3}{\sigma^3}, \qquad \gamma_4 = \frac{E(R_t - \mu)^4}{\sigma^4} \qquad \text{with } \sigma^2 = 1 + \frac{1}{2}\text{SR}^2. \tag{2.15}$$

Christie [9] derived the asymptotic distribution of the Sharpe ratio under the more relaxed assumption of stationarity and ergodicity. Opdyke [4] interestingly showed that the derivation provided by Christie [9] under the non-IID returns condition was in fact identical to the one provided by Mertens [8]. For a complete discussion and proof see Opdyke [4]. For our purposes, we will use asymptotic distributions (2.12) and (2.14) to compare to our proposed likelihood-based approach. In the next section below we review the likelihood-based approach to inference and in the following section we apply the methodology to the Sharpe ratio.

## 3. Likelihood-Based Approach to Inference

In this section we review standard first-order likelihood methods and then turn our attention to higher-order methods. For this section, the underlying model under consideration is as follows: $Y = (Y_1, \ldots, Y_n)$ is a vector of IID random variables with density given by $f(\cdot; \theta)$, where $\theta$ is a $k$-dimensional vector of parameters. Statistical inference concerns drawing conclusions about $\theta$ or a function of $\theta$, namely, $\psi = \psi(\theta)$, given an observed sample $y = (y_1, \ldots, y_n)$ and in this paper we assume that the interest parameter is $\psi(\theta)$. We further assume that this interest parameter is scalar. Let $\lambda = \lambda(\theta)$ be a $(k-1)$-dimensional nuisance parameter vector. The log-likelihood function of $\theta$ is

$$l(\theta) = l(\theta; y) = \log\left(\prod_{i=1}^n f(y_i; \theta)\right). \tag{3.1}$$

Likelihood analysis typically involves the maximum likelihood estimator (MLE) $\hat{\theta} = (\hat{\psi}, \hat{\lambda}')' = \arg\max_\theta l(\theta; y)$ and the constrained maximum likelihood estimator $\tilde{\theta} = (\psi, \hat{\lambda}')' = \arg\max_\lambda l(\theta; y)$, where the maximum is taken for fixed values of $\psi$. The constrained MLE can be solved by maximizing $l(\theta)$ subject to $\psi(\theta) = \psi$. This is often done using the Lagrange multiplier method where the Lagrangean is given by the following:

$$H(\theta, \alpha) = l(\theta; y) + \alpha[\psi(\theta) - \psi]. \tag{3.2}$$

The Lagrange multiplier is given by $\hat{\alpha}$. The tilted log-likelihood is defined as follows:

$$\tilde{l}(\theta) = l(\theta) + \hat{\alpha}[\psi(\theta) - \psi]. \tag{3.3}$$

Using the above information, standard first-order departure measures can be obtained. These measures include the Wald departure (the standardized maximum likelihood estimator) and the signed square root log-likelihood ratio statistic:

$$q \equiv q(\psi) = (\widehat{\psi} - \psi)\left[\mathrm{var}(\widehat{\psi})\right]^{-1/2}, \tag{3.4}$$

$$R \equiv R(\psi) = \mathrm{sgn}(\widehat{\psi} - \psi)\left\{2\left[l\left(\widehat{\theta}; y\right) - l\left(\widetilde{\theta}; y\right)\right]\right\}^{1/2}. \tag{3.5}$$

Note that $\mathrm{var}(\widehat{\psi})$ can be estimated by using the Delta method

$$\widehat{\mathrm{var}}(\widehat{\psi}) \approx \psi'_\theta\left(\widehat{\theta}\right)\left[\mathrm{var}\left(\widehat{\theta}\right)\right]\psi_\theta\left(\widehat{\theta}\right), \tag{3.6}$$

where the variance of $\widehat{\theta}$, $\mathrm{var}(\widehat{\theta})$, can be estimated by using either the Fisher expected information matrix

$$E\left[-\frac{\partial^2 l(\theta)}{\partial\theta\partial\theta'}\right], \tag{3.7}$$

or the observed information matrix evaluated at the MLE

$$j_{\theta\theta'}\left(\widehat{\theta}\right) = -\left.\frac{\partial^2 l(\theta)}{\partial\theta\partial\theta'}\right|_{\theta=\widehat{\theta}}. \tag{3.8}$$

The latter is preferred however because of the simplicity in calculation. The $p$ value function for these methods are $p(\psi) = \Phi(q)$ and $p(\psi) = \Phi(R)$, where $\Phi$ represents the standard normal cumulative distribution function. These $p$ value functions have order of convergence $O(n^{-1/2})$ and are correspondingly referred to as first-order methods (note that (3.5) is invariant to reparameterization whereas (3.4) is not. However, in practice, (3.4) is preferred because of its simplicity of use. Doganaksoy and Schmee [10] illustrate that (3.5) has better coverage than (3.4) in the cases they examined). A $100(1-\alpha)\%$ confidence interval is then given by the following:

$$\left\{\min\left[p^{-1}\left(\frac{\alpha}{2}\right), p^{-1}\left(1-\frac{\alpha}{2}\right)\right], \max\left[p^{-1}\left(\frac{\alpha}{2}\right), p^{-1}\left(1-\frac{\alpha}{2}\right)\right]\right\}. \tag{3.9}$$

Approximations aimed at improving the accuracy of the first-order methods have been worked on during the past three decades. Higher-order asymptotic approximations for the distribution of an estimate can be constructed using expansion methods of which Edgeworth expansions and saddlepoint approximations are most common. Edgeworth approximations tend to produce good approximations near the center of a distribution, with relatively poorer tail approximations, while saddlepoint approximations produce remarkably accurate tail approximations. Given we are interested in tail probabilities, our focus will thus be on approximation methods that use the saddlepoint method. We will not review the saddlepoint approximation literature here but the interested reader is directed to Daniels [11, 12] and Barndorff-Nielsen and Cox [13].

For a canonical exponential family model with canonical parameter vector $\theta = (\psi, \lambda')'$ and density

$$f(y; \theta) = \exp\{\psi u(y) + \lambda' v(y) - c(\theta) + d(y)\}, \tag{3.10}$$

where $(U(Y), V(Y))$ is a minimal sufficient statistic for $\theta$, the saddlepoint-based approximation for the density of $\hat{\theta}$ is given as follows:

$$f\left(\hat{\theta}; \theta\right) \doteq c \left| j_{\theta\theta'}\left(\hat{\theta}\right) \right|^{1/2} \exp\left[l(\theta; y) - l\left(\hat{\theta}; y\right)\right], \tag{3.11}$$

where $c$ is a normalizing constant. This extremely important approximation to the density of the maximum likelihood estimator is referred to as Barndorff-Nielsen's $p^*$ formula and can be found in Barndorff-Nielsen [14] (this approximation is valid for more general models but requires the existence of an ancillary statistic). Thus the marginal density for $\hat{\psi}$ can be obtained. In general however, the closed form of the marginal density is not available.

Fraser et al. [15] showed that the result by Lugannani and Rice [16] can be applied to approximate the $p$ value function and it takes the form:

$$p(\psi) \doteq \Phi(R) + \phi(R) \left(\frac{1}{R} - \frac{1}{Q}\right). \tag{3.12}$$

The statistic $Q$ is given by the following:

$$Q \equiv Q(\psi) = (\hat{\psi} - \psi) \left| j_{\theta\theta'}\left(\hat{\theta}\right) \right|^{1/2} \left| j_{\lambda\lambda'}\left(\tilde{\theta}\right) \right|^{-1/2}, \tag{3.13}$$

where $j_{\theta\theta'}(\hat{\theta})$ is the observed information matrix and $j_{\lambda\lambda'}(\tilde{\theta})$ is the observed nuisance information matrix defined as follows:

$$j_{\theta\theta'}\left(\hat{\theta}\right) = -\frac{\partial^2 l(\theta; y)}{\partial\theta\partial\theta'}\bigg|_{\theta=\hat{\theta}}, \tag{3.14}$$

$$j_{\lambda\lambda'}\left(\tilde{\theta}\right) = -\frac{\partial^2 l(\theta; y)}{\partial\lambda\partial\lambda'}\bigg|_{\theta=\tilde{\theta}}. \tag{3.15}$$

The term $|j_{\theta\theta'}(\hat{\theta})||j_{\lambda\lambda'}(\tilde{\theta})|^{-1}$ can be viewed as an estimate of the variance of $\hat{\psi}$, where the elimination of the nuisance parameter $\lambda$ has been taken into consideration. Note that $Q$ is a standardized maximum likelihood departure in the canonical parameter scale. The statistic $R$ is the log-likelihood ratio given by (3.5). An alternate approximation to (3.12) is given by Barndorff-Nielsen [17]:

$$p(\psi) \doteq \Phi\left(R - \frac{1}{R}\log\left(\frac{R}{Q}\right)\right) \doteq \Phi(R^*). \tag{3.16}$$

These approximations are asymptotically equivalent (see [18]) and have order of convergence $O(n^{-3/2})$. They are correspondingly referred to as third-order methods. Hence $100(1 - \alpha)\%$ confidence interval for $\psi$ can be obtained from (3.9).

Consider now a general full rank exponential family model with parameter vector $\theta$ and density given by the following:

$$f(y;\theta) = \exp\{\varphi'(\theta)t(y) - c(\theta) + h(t(y))\}, \tag{3.17}$$

where $\varphi(\theta)$ is the canonical parameter and $t(y)$ is the canonical variable. In order to use either the Lugannani and Rice approximation given in (3.12) or the Barndorff-Nielsen approximation given in (3.16) for inference on our interest parameter $\psi(\theta)$, we need to calculate both $R$ and $Q$ in $\varphi(\theta)$ scale. Note that $R$ calculated in the original $\theta$ scale is equivalent to $R$ calculated in $\varphi(\theta)$ scale. Moreover, Fraser and Reid [19] derive a $Q$ in the $\varphi(\theta)$ scale. Their methodology involves replacing the parameter of interest $\psi$ by a linear function of the $\varphi(\theta)$ coordinates (for a detailed discussion and derivation see Fraser and Reid [19]). This newly calibrated parameter is given by the following:

$$\chi(\theta) = \psi_{\theta'}(\tilde{\theta})\varphi_{\theta'}^{-1}(\tilde{\theta})\varphi(\theta). \tag{3.18}$$

Then $|\chi(\hat{\theta}) - \chi(\tilde{\theta})|$ measures the departure of $|\hat{\psi} - \psi|$ in the $\varphi(\theta)$ scale. The function $Q$ is then constructed as follows:

$$Q = \text{sgn}\left(\psi\left(\hat{\theta}\right) - \psi\right)\left|\chi\left(\hat{\theta}\right) - \chi\left(\tilde{\theta}\right)\right|\left[\widehat{\text{var}}\left(\chi(\hat{\theta}) - \chi(\tilde{\theta})\right)\right]^{-1/2}. \tag{3.19}$$

Fraser and Reid [19] show that

$$\widehat{\text{var}}\left(\chi\left(\hat{\theta}\right) - \chi\left(\tilde{\theta}\right)\right) \approx \frac{\psi'_{\theta'}\left(\tilde{\theta}\right)\tilde{j}_{\theta\theta'}^{-1}\left(\tilde{\theta}\right)\psi_{\theta'}\left(\tilde{\theta}\right)\left|\tilde{j}_{\theta\theta'}\left(\tilde{\theta}\right)\right|\left|\varphi_{\theta'}\left(\tilde{\theta}\right)\right|^{-2}}{\left|j_{\theta\theta'}\left(\hat{\theta}\right)\right|\left|\varphi_{\theta'}\left(\hat{\theta}\right)\right|^{-2}}, \tag{3.20}$$

where $j_{\theta\theta'}(\hat{\theta})$ is defined in (3.14) and $\tilde{j}_{\theta\theta'}(\tilde{\theta})$ is defined as follows:

$$\tilde{j}_{\theta\theta'}\left(\tilde{\theta}\right) = -\left.\frac{\partial^2 \tilde{l}(\theta)}{\partial\theta\partial\theta'}\right|_{\theta=\tilde{\theta}}, \tag{3.21}$$

where $\tilde{l}(\theta)$ is the tilted log-likelihood function defined in (3.3). The function $Q$ in (3.19) along with $R$ is used in the Lugannani and Rice approximation in (3.12) or the Barndorff-Nielsen expression in (3.16) which implicitly yields a new $R^*$.

In the next section we apply the above methodology and use Barndorff-Nielsen's approximation to obtain highly accurate $p$ values to test particular hypothesized values of the Sharpe ratio. We further note that for all our examples and simulations in this paper, the Lugannani and Rice expression given in (3.12) was also used but we do not report results from this method as they are virtually identical to those produced by the Barndorff-Nielsen expression.

## 4. Methodology for the Sharpe Ratio

We emphasize that while the third-order methodology discussed in the previous section is applicable under any parametric distributional assumptions, for expositional clarity we demonstrate the use of the method under the assumption of IID normal returns (see Fraser et al. [20] for the general model setup). Consider a fund with return at time $t$ given by $R_t$, $(t = 1, \ldots, T)$. And let $R_f$ be the mean return for the risk-free asset. We assume that $R_1, \ldots R_T$ IID $N(\mu, \sigma^2)$. The interest parameter is the Sharpe Ratio:

$$\text{SR} = \frac{\mu - R_f}{\sigma} = \psi, \tag{4.1}$$

and $\theta = (\mu, \sigma^2)'$. The log-likelihood and related first and second derivatives, denoted by $l.(\theta)$ and $l..(\theta)$, respectively, are given as follows:

$$l(\theta) = l\left(\mu, \sigma^2\right) = -\frac{T}{2}\log \sigma^2 - \frac{1}{2\sigma^2}\Sigma(R_t - \mu)^2,$$

$$l_\mu(\theta) = \frac{1}{\sigma^2}\Sigma(R_t - \mu),$$

$$l_{\sigma^2}(\theta) = -\frac{T}{2\sigma^2} + \frac{1}{2\sigma^4}\Sigma(R_t - \mu)^2,$$

$$l_{\mu\mu}(\theta) = -\frac{T}{\sigma^2}, \tag{4.2}$$

$$l_{\mu\sigma^2}(\theta) = -\frac{1}{\sigma^4}\Sigma(R_t - \mu),$$

$$l_{\sigma^2\sigma^2}(\theta) = \frac{T}{2\sigma^4} - \frac{1}{\sigma^6}\Sigma(R_t - \mu)^2.$$

To obtain the maximum likelihood estimators, the first order conditions

$$l_\mu(\theta) = 0,$$

$$l_{\sigma^2}(\theta) = 0 \tag{4.3}$$

are solved, and we have

$$\hat{\mu} = \overline{R},$$

$$\hat{\sigma}^2 = \frac{\Sigma\left(R_t - \overline{R}\right)^2}{T}. \tag{4.4}$$

The estimated Sharpe ratio is then

$$\widehat{\text{SR}} = \frac{\hat{\mu} - R_f}{\hat{\sigma}}. \tag{4.5}$$

The observed information matrix is given as $j_{\theta\theta'}(\theta) = -l_{\theta\theta'}(\theta)$ which can be expressed as follows:

$$j_{\theta\theta'}(\theta) = \begin{bmatrix} -l_{\mu\mu}(\theta) & -l_{\mu\sigma^2}(\theta) \\ -l_{\sigma^2\mu}(\theta) & -l_{\sigma^2\sigma^2}(\theta) \end{bmatrix} = \begin{bmatrix} \dfrac{T}{\sigma^2} & \dfrac{1}{\sigma^4}\Sigma(R_t - \mu) \\ \dfrac{1}{\sigma^4}\Sigma(R_t - \mu) & -\dfrac{T}{2\sigma^4} + \dfrac{1}{\sigma^6}\Sigma(R_t - \mu)^2 \end{bmatrix}. \tag{4.6}$$

The observed information matrix is then evaluated at the MLE:

$$j_{\theta\theta'}(\hat{\theta}) = \begin{pmatrix} \dfrac{T}{\hat{\sigma}^2} & 0 \\ 0 & \dfrac{T}{2\hat{\sigma}^2} \end{pmatrix} \tag{4.7}$$

with determinant equal to

$$\left| j_{\theta\theta'}(\hat{\theta}) \right| = \left( \dfrac{T}{\hat{\sigma}^2} \right)\left( \dfrac{T}{2\hat{\sigma}^4} \right) = \dfrac{T^2}{2\hat{\sigma}^6}. \tag{4.8}$$

To solve for the constrained MLE, the Lagrange multiplier method from (3.2) is used

$$H(\theta, \alpha) = l(\theta) + \alpha\left[ \dfrac{\mu - R_f}{\sigma} - \psi \right], \tag{4.9}$$

with first derivatives equal to

$$H_\mu(\theta, \alpha) = l_\mu(\theta) + \dfrac{\alpha}{\sigma},$$

$$H_{\sigma^2}(\theta, \alpha) = l_{\sigma^2}(\theta) - \dfrac{\alpha(\mu - R_f)}{2\sigma^3}, \tag{4.10}$$

$$H_\alpha(\theta, \alpha) = \dfrac{\mu - R_f}{\sigma} - \psi.$$

Setting these first derivatives equal to zero, we have:

$$\hat{\alpha} = -\tilde{\sigma}l_\mu(\tilde{\theta}), \tag{4.11}$$

$$\tilde{\mu} = R_f + \psi\tilde{\sigma}, \tag{4.12}$$

$$l_{\sigma^2}(\tilde{\theta}) + \tilde{\sigma}l_\mu(\tilde{\theta})\dfrac{\tilde{\mu} - R_f}{2\tilde{\sigma}^3} = 0. \tag{4.13}$$

Working on (4.13) we have

$$-\dfrac{T}{2\tilde{\sigma}^2} + \dfrac{1}{2\tilde{\sigma}^4}\Sigma(R_t - \tilde{\mu})^2 + \tilde{\sigma}\dfrac{1}{\tilde{\sigma}^2}\Sigma(R_t - \tilde{\mu})\dfrac{\tilde{\mu} - R_f}{2\tilde{\sigma}^3} = 0 \tag{4.14}$$

which further simplifies to

$$-T\tilde{\sigma}^2 - \psi\Sigma(R_t - R_f)\tilde{\sigma} + \Sigma(R_t - R_f)^2 = 0. \tag{4.15}$$

If we let $r_t = R_t - R_f$ and $\bar{r} = \Sigma r_t/T$, then we have

$$\tilde{\sigma}^2 + \psi\bar{r}\tilde{\sigma} - \frac{\Sigma r_t^2}{T} = 0 \tag{4.16}$$

with

$$\tilde{\sigma} = \frac{-\psi\bar{r} + \sqrt{(\psi\bar{r})^2 + 4(\Sigma r_t^2/T)}}{2}, \quad \text{since } \tilde{\sigma} > 0. \tag{4.17}$$

The constrained MLE are therefore given by:

$$\tilde{\sigma} = \frac{-\psi\bar{r} + \sqrt{(\psi\bar{r})^2 + 4(\Sigma r_t^2/T)}}{2},$$

$$\tilde{\mu} = R_f + \psi\tilde{\sigma}, \tag{4.18}$$

$$\hat{\alpha} = T\left(\psi - \frac{\bar{r}}{\tilde{\sigma}}\right),$$

so that $\tilde{\theta} = (\tilde{\mu}, \tilde{\sigma}^2)'$. Using these constrained maximum likelihood estimators we are able to obtain the titled log-likelihood as defined in (3.3), and the corresponding first and second derivatives are:

$$\tilde{l}(\theta) = l(\theta) + \hat{\alpha}\left(\frac{\mu - R_f}{\sigma} - \psi\right),$$

$$\tilde{l}_\mu(\theta) = l_\mu(\theta) + \frac{\hat{\alpha}}{\sigma},$$

$$\tilde{l}_{\sigma^2}(\theta) = l_{\sigma^2}(\theta) - \hat{\alpha}\left(\frac{\mu - R_f}{2\sigma^3}\right),$$

$$\tilde{l}_{\mu\mu}(\theta) = l_{\mu\mu}(\theta), \tag{4.19}$$

$$\tilde{l}_{\mu\sigma^2}(\theta) = l_{\mu\sigma^2}(\theta) - \frac{\hat{\alpha}}{2\sigma^3},$$

$$\tilde{l}_{\sigma^2\sigma^2}(\theta) = l_{\sigma^2\sigma^2}(\theta) + \frac{3}{4}\frac{\hat{\alpha}(\mu - R_f)}{\sigma^5}.$$

We can use these second derivatives to calculate $\tilde{j}_{\theta\theta'}(\tilde{\theta})$ as

$$\tilde{j}_{\theta\theta'}(\tilde{\theta}) = \begin{pmatrix} -\tilde{l}_{\mu\mu}(\tilde{\theta}) & -\tilde{l}_{\mu\sigma^2}(\tilde{\theta}) \\ -\tilde{l}_{\sigma^2\mu}(\tilde{\theta}) & -\tilde{l}_{\sigma^2\sigma^2}(\tilde{\theta}) \end{pmatrix}. \tag{4.20}$$

For the normal model, the canonical parameter is given by:

$$\varphi(\theta) = \left( \frac{\mu}{\sigma^2}, \frac{1}{\sigma^2} \right)'. \tag{4.21}$$

Two related matrices derived from the canonical parameter are given below

$$\varphi_\theta(\theta) = \begin{pmatrix} \dfrac{1}{\sigma^2} & -\dfrac{\mu}{\sigma^4} \\ 0 & -\dfrac{1}{\sigma^4} \end{pmatrix}, \tag{4.22}$$

$$\varphi_\theta^{-1}(\theta) = \begin{pmatrix} \sigma^2 & -\mu\sigma^2 \\ 0 & -\sigma^4 \end{pmatrix}. \tag{4.23}$$

With our parameter of interest

$$\psi(\theta) = \frac{\mu - R_f}{\sigma} \tag{4.24}$$

we have

$$\psi_\theta(\theta) = \left( \psi_\mu(\theta), \psi_{\sigma^2}(\theta) \right) = \left( \frac{1}{\sigma}, -\frac{\mu - R_f}{2\sigma^3} \right). \tag{4.25}$$

Quantities (4.21), (4.23), and (4.25) can be used to calculate (3.18). The quantity $\widehat{\mathrm{var}}(\chi(\hat{\theta}) - \chi(\tilde{\theta}))$ can further be calculated. We now have all the ingredients to calculate $R^*(\psi)$ given in (3.16).

## 5. Examples and Simulations

In this section, we provide examples and simulations for inference on the Sharpe ratio. We compute confidence intervals and $p$ values using our proposed third-order method given in (3.16). We label this method "proposed" in the tables below. We compare our results to those obtained by using the existing methods, in particular, using both the Jobson and Korkie [5] asymptotic distribution in (2.12) and the Mertens [8] distribution given in (2.14). These methods are labeled as "Jobson and Korkie" and "Mertens," respectively. Results from the signed square root log-likelihood ratio statistic in (3.5) are additionally provided and labeled "likelihood ratio."

**Table 1:** Monthly return prices.

| Month | Fund | Market | Cash |
|---|---|---|---|
| 1 | 1.155809268 | 1.021191715 | 1.071489631 |
| 2 | 1.165211551 | 1.022641807 | 1.07422193 |
| 3 | 1.183344526 | 1.060673856 | 1.077202896 |
| 4 | 1.184016118 | 1.05100051 | 1.080281901 |
| 5 | 1.186030893 | 1.059513614 | 1.083405716 |
| 6 | 1.163196776 | 1.041851522 | 1.08684553 |
| 7 | 1.190732035 | 1.081254347 | 1.090341549 |
| 8 | 1.19677636 | 1.081621973 | 1.093957849 |
| 9 | 1.216252518 | 1.110263323 | 1.097941679 |
| 10 | 1.222296843 | 1.113272137 | 1.102068109 |
| 11 | 1.224311618 | 1.127132375 | 1.106210049 |
| 12 | 1.234385376 | 1.141573195 | 1.110450484 |

**Table 2:** 95% Confidence Intervals for the sharpe Ratio.

| Method | 95% CI for SR (Fund) | 95% CI for SR (Market) |
|---|---|---|
| Jobson and Korkie | (−0.2238, 0.9441) | (−0.1009, 1.0992) |
| Mertens | (−0.2838, 1.0042) | (−0.0644, 1.0626) |
| Likelihood Ratio | (−0.2087, 0.9624) | (−0.0800, 1.1261) |
| Proposed | (−0.2319, 0.9383) | (−0.1122, 1.0924) |

### 5.1. Examples

The dataset for our examples consists of monthly return prices for three time series. The first series represents return prices for a large-cap mutual fund (Fund), the second for a market index (Market), and the third for 90-day Treasury bills (Cash). This data spans a period of one year. The data is listed in Table 1 and originates from the Matlab Financial Toolbox User's Guide (data in the User's Guide is provided for a 5-year period. We focus on the most recent year of data. See Section 4: Investment Performance Metrics of the manual for further information).

Table 2 reports 95% confidence intervals for the Sharpe ratio separately for the large-cap mutual fund and the market index for the four methods discussed previously. Monthly returns are calculated from the return prices listed in Table 1 and are used to calculate the Sharpe ratio. The mean of the 90-day Treasury returns is used as the risk-free rate. Table 2 shows that the confidence intervals obtained from the four methods produce rather different results. Theoretically, the proposed method has third-order accuracy whereas the remaining three methods do not. Although the confidence intervals produced by the Jobson and Korkie [5] approximation have the best concordance with those produced by the proposed method, the difference is still noticeable. This result will be borne out in the simulations as well.

The $p$ value functions calculated from the methods discussed in this paper for the Sharpe ratio for the large-cap mutual fund and the Sharpe ratio for the market index are plotted in Figures 1 and 2, respectively. These significance functions can be used to obtain $p$ values for specific hypothesized values of the Sharpe ratio. As we are typically interested in tail probabilities which tend to be small, it is important to estimate such probabilities with precision. A few values with their corresponding $p$ values are provided in Tables 3 and 4 for the market fund and market index, respectively. From these tables we can see that the $p$ values
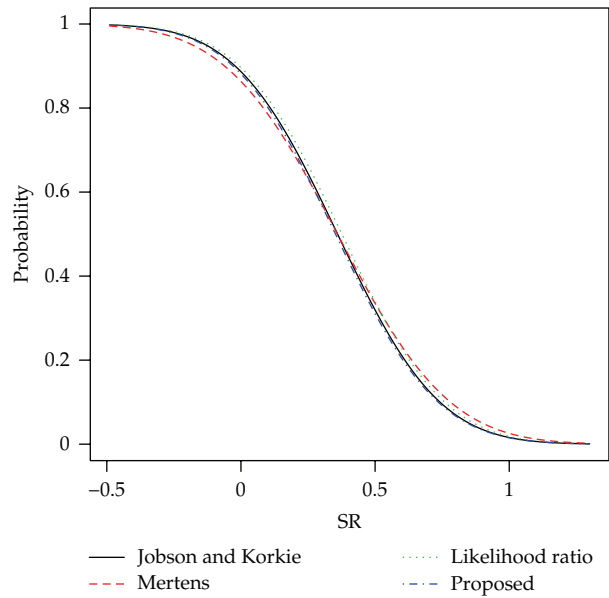
**Figure 1:** $p$ value function for Fund.



**Figure 2:** $p$ value function for Market.
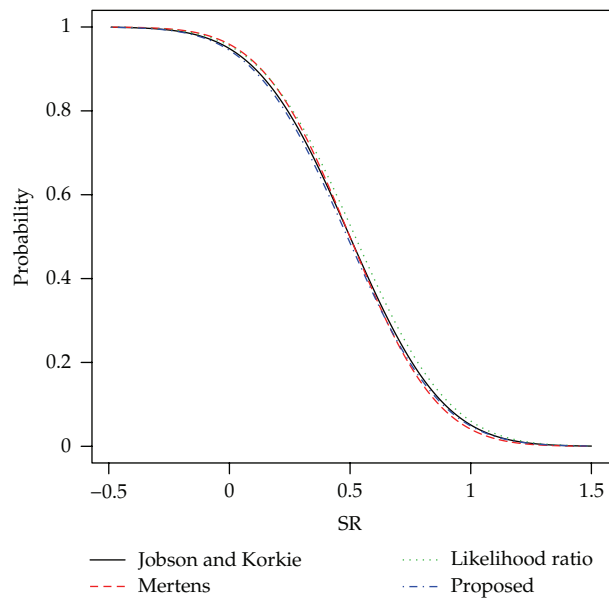
vary across the methods. If for instance, interest is on testing whether SR = 1 for either fund, then the corresponding $p$ values for such a test are given in the last column of Tables 3 and 4. Focussing on the market index and using a 5% level of significance, the Sharpe ratio may or may not be statistically significant depending on the method chosen for the hypothesis test.

**Table 3:** Fund: $p$ values for SR.

| Method | SR | | | | | |
|---|---|---|---|---|---|---|
| | −0.25 | 0 | 0.25 | 0.50 | 0.75 | 1 |
| Jobson and Korkie | 0.9797 | 0.8867 | 0.6442 | 0.3194 | 0.0953 | 0.0159 |
| Mertens | 0.9684 | 0.8635 | 0.6313 | 0.3352 | 0.1177 | 0.0257 |
| Likelihood ratio | 0.9821 | 0.8962 | 0.6637 | 0.3393 | 0.1055 | 0.0185 |
| Proposed | 0.9784 | 0.8813 | 0.6344 | 0.3107 | 0.0917 | 0.0151 |

**Table 4:** Market: $p$ values for SR.

| Method | SR | | | | | |
|---|---|---|---|---|---|---|
| | −0.25 | 0 | 0.25 | 0.50 | 0.75 | 1 |
| Jobson and Korkie | 0.9928 | 0.9485 | 0.7921 | 0.4989 | 0.2063 | 0.0509 |
| Mertens | 0.9954 | 0.9588 | 0.8070 | 0.4988 | 0.1914 | 0.0407 |
| Likelihood ratio | 0.9941 | 0.9553 | 0.8114 | 0.5277 | 0.2289 | 0.0603 |
| Proposed | 0.9921 | 0.9444 | 0.7813 | 0.4849 | 0.1976 | 0.0484 |

### 5.2. Simulations

Two simulation studies of size 10,000 were performed to compare the three existing methods to the proposed third-order method. The first simulation was constructed to mimic the mutual fund return data from the above example and the second to mimic the market index returns: $X \sim N(0.007378393, 0.01137605)$, $Y \sim N(0.01211267, 0.01789324)$ and $R_f = 0.003181247$. 90%, 95%, and 99% confidence intervals for the Sharpe ratio were obtained for each sample and the lower probability error, the upper probability error, and the central coverage were recorded. Lower (upper) probability error refers to the proportion of times the true parameter value falls below (above) the confidence interval. Central coverage is defined as the proportion of times the true parameter values falls within the confidence intervals. Tables 5 and 6 report the results from these simulations for the fund and market returns, respectively. For reference, the nominal values and the corresponding standard errors have been included for each the lower probability error, the upper probability error, and the central coverage.

From these results tables it is clear that the third-order method outperforms the other three methods based on the criteria we examined. It must be noted however, that the Jobson and Korkie [5] method provides surprisingly good results. We note that other values for the parameters of the simulation were chosen (but not reported) and the results were consistent with the reported results. Standard errors for the simulations can easily be computed. This information has been included in the tables and labeled "standard error." From these standard errors, it can be seen that the proposed method produces results that are uniformly within three standard deviations of the nominal value. The other three methods produce less satisfactory results.

## 6. The Two-Sample Case

Suppose one is interested in testing hypotheses concerning the Sharpe ratios of two funds $X$ and $Y$. For instance, one may be interested in testing the null hypothesis $SR_X \geq SR_Y$ or $SR_X = SR_Y$ against the alternative hypothesis $SR_X < SR_Y$ or $SR_X \neq SR_Y$, respectively.

**Table 5:** Results for simulation study for fund for $n = 12$.

| CI | Method | Lower error | Upper error | Central coverage |
|---|---|---|---|---|
| | Jobson and Korkie | 0.0586 | 0.0521 | 0.8893 |
| | Mertens | 0.0778 | 0.0507 | 0.9715 |
| 90% | Likelihood ratio | 0.0712 | 0.0529 | 0.8759 |
| | Proposed | 0.0512 | 0.0502 | 0.8986 |
| | *Nominal* | 0.0500 | 0.0500 | 0.9000 |
| | *Standard error* | 0.0022 | 0.0022 | 0.0030 |
| | Jobson and Korkie | 0.0319 | 0.0267 | 0.9414 |
| | Mertens | 0.0514 | 0.0278 | 0.9208 |
| 95% | Likelihood ratio | 0.0416 | 0.0285 | 0.9299 |
| | Proposed | 0.0279 | 0.0251 | 0.9470 |
| | *Nominal* | 0.0250 | 0.0250 | 0.9500 |
| | *Standard error* | 0.0016 | 0.0016 | 0.0022 |
| | Jobson and Korkie | 0.0059 | 0.0065 | 0.9876 |
| | Mertens | 0.0208 | 0.0072 | 0.9720 |
| 99% | Likelihood ratio | 0.0098 | 0.0072 | 0.9830 |
| | Proposed | 0.0048 | 0.0052 | 0.9900 |
| | *Nominal* | 0.0050 | 0.0050 | 0.9900 |
| | *Standard error* | 0.0007 | 0.0007 | 0.0010 |

**Table 6:** Results for simulation study for market for $n = 12$.

| CI | Method | Lower error | Upper error | Central coverage |
|---|---|---|---|---|
| | Jobson and Korkie | 0.0622 | 0.0518 | 0.8860 |
| | Mertens | 0.0930 | 0.0514 | 0.8556 |
| 90% | Likelihood ratio | 0.0821 | 0.0513 | 0.8666 |
| | Proposed | 0.0528 | 0.0514 | 0.8958 |
| | *Nominal* | 0.0500 | 0.0500 | 0.9000 |
| | *Standard error* | 0.0022 | 0.0022 | 0.0030 |
| | Jobson and Korkie | 0.0320 | 0.0290 | 0.9390 |
| | Mertens | 0.0500 | 0.0284 | 0.9216 |
| 95% | Likelihood ratio | 0.0407 | 0.0303 | 0.9290 |
| | Proposed | 0.0280 | 0.0275 | 0.9445 |
| | *Nominal* | 0.0250 | 0.0250 | 0.9500 |
| | *Standard error* | 0.0016 | 0.0016 | 0.0022 |
| | Jobson and Korkie | 0.0053 | 0.0061 | 0.9886 |
| | Mertens | 0.0245 | 0.0073 | 0.9682 |
| 99% | Likelihood ratio | 0.0100 | 0.0069 | 0.9831 |
| | Proposed | 0.0046 | 0.0053 | 0.9901 |
| | *Nominal* | 0.0050 | 0.0050 | 0.9900 |
| | *Standard error* | 0.0007 | 0.0007 | 0.0010 |

In this section we apply the third-order methodology described in Section 3 to test the difference between the Sharpe ratios of two funds. Consider two funds with returns given by $(x_1, \ldots, x_n)$ and $(y_1, \ldots, y_m)$. Further assume for expositional clarity that these returns are identically and independently distributed as $N(\mu_X, \sigma_X^2)$ and $N(\mu_Y, \sigma_Y^2)$, respectively.

The mean return for the risk-free asset is given by $R_f$. Our interest is in testing the difference in Sharpe ratios of the two funds. The interest parameter $\psi$ is then defined as follows:

$$\psi = \frac{\mu_X - R_f}{\sigma_X} - \frac{\mu_Y - R_f}{\sigma_Y}. \tag{6.1}$$

The log-likelihood function for this problem is written as follows:

$$l(\theta) = l\left(\mu_X, \sigma_X^2, \mu_Y, \sigma_Y^2\right) = -\frac{n}{2}\log\sigma_X^2 - \frac{1}{2\sigma_X^2}\Sigma(x_i - \mu_X)^2 - \frac{m}{2}\log\sigma_Y^2 - \frac{1}{2\sigma_Y^2}\Sigma(y_i - \mu_Y)^2. \tag{6.2}$$

The corresponding first derivatives are

$$l_{\mu_X} = \frac{1}{\sigma_X^2}\Sigma(x_i - \mu_X),$$

$$l_{\sigma_X^2} = -\frac{n}{2\sigma_X^2} + \frac{1}{2\sigma_X^4}\Sigma(x_i - \mu_X)^2,$$

$$l_{\mu_Y} = \frac{1}{\sigma_Y^2}\Sigma(y_j - \mu_Y), \tag{6.3}$$

$$l_{\sigma_Y^2} = -\frac{m}{2\sigma_Y^2} + \frac{1}{2\sigma_Y^4}\Sigma(y_j - \mu_Y)^2.$$

These first derivatives are set equal to zero and simultaneously solved for the overall MLE:

$$\widehat{\mu}_X = \overline{x},$$

$$\widehat{\sigma}_X^2 = \frac{\Sigma(x_i - \overline{x})^2}{n},$$

$$\widehat{\mu}_Y = \overline{y}, \tag{6.4}$$

$$\widehat{\sigma}_Y^2 = \frac{\Sigma(y_j - \overline{y})^2}{m}.$$

The second derivatives are required to obtain the observed information matrix:

$$
j_{\theta\theta}(\theta) =
\begin{bmatrix}
\dfrac{n}{\sigma_X^2} & \dfrac{1}{\sigma_X^4}\Sigma(x_i - \mu_X) & 0 & 0 \\[2ex]
\dfrac{1}{\sigma_X^4}\Sigma(x_i - \mu_X) & -\dfrac{n}{2\sigma_X^4} + \dfrac{1}{\sigma_X^6}\Sigma(x_i - \mu_X)^2 & 0 & 0 \\[2ex]
0 & 0 & \dfrac{m}{\sigma_Y^2} & \dfrac{1}{\sigma_Y^4}\Sigma(y_j - \mu_Y) \\[2ex]
0 & 0 & \dfrac{1}{\sigma_Y^4}\Sigma(y_j - \mu_Y) & -\dfrac{m}{2\sigma_Y^4} + \dfrac{1}{\sigma_Y^6}\Sigma(y_j - \mu_Y)^2
\end{bmatrix}.
\tag{6.5}
$$

Evaluating this observed information matrix at the MLE produces $j_{\theta\theta}(\widehat{\theta})$ with corresponding determinant denoted by $|j_{\theta\theta}(\widehat{\theta})|$.

To solve for the constrained MLE, the Lagrange multiplier method is used

$$
H(\theta) = l(\theta) + \alpha\left[\frac{\mu_X - R_f}{\sqrt{\sigma_X^2}} - \frac{\mu_Y - R_f}{\sqrt{\sigma_Y^2}} - \psi\right],
\tag{6.6}
$$

with the following first derivatives:

$$
H_{\mu_X} = \frac{1}{\sigma_X^2}\Sigma(x_i - \mu_X) + \frac{\alpha}{\sqrt{\sigma_X^2}},
$$

$$
H_{\sigma_X^2} = -\frac{n}{2\sigma_X^2} + \frac{1}{2\sigma_X^4}\Sigma(x_i - \mu_X)^2 - \frac{\alpha(\mu_X - R_f)}{2(\sigma_X^2)^{3/2}},
$$

$$
H_{\mu_Y} = \frac{1}{\sigma_Y^2}\Sigma(y_j - \mu_Y) - \frac{\alpha}{\sqrt{\sigma_Y^2}},
\tag{6.7}
$$

$$
H_{\sigma_X^2} = -\frac{m}{2\sigma_Y^2} + \frac{1}{2\sigma_Y^4}\Sigma(y_j - \mu_Y)^2 + \frac{\alpha(\mu_Y - R_f)}{2(\sigma_Y^2)^{3/2}},
$$

$$
H_\alpha = \frac{\mu_X - R_f}{\sqrt{\sigma_X^2}} - \frac{\mu_Y - R_f}{\sqrt{\sigma_Y^2}} - \psi.
$$

Setting these derivatives equal to zero and solving the resulting system produces the constrained MLE. The constrained MLE can be obtained by solving the following iteratively:

$$
\tilde{\sigma}_X^2 = \frac{\Sigma(x_i - \tilde{\mu}_X)^2 + (\tilde{\mu}_X - R_f)\Sigma(x_i - \tilde{\mu}_X)}{n},
$$

$$
\tilde{\sigma}_Y^2 = \frac{\Sigma(y_j - \tilde{\mu}_Y)^2 + (\tilde{\mu}_Y - R_f)\Sigma(y_j - \tilde{\mu}_Y)}{m},
$$

$$
\tilde{\mu}_X = \frac{m\tilde{\sigma}_Y R_f + n\overline{x}\tilde{\sigma}_Y + m\overline{y}\tilde{\sigma}_X - m\tilde{\sigma}_X R_f + m\psi\tilde{\sigma}_X\tilde{\sigma}_Y}{(n+m)\tilde{\sigma}_X},
$$

$$
\tilde{\mu}_Y = \frac{n\overline{x}\tilde{\sigma}_Y - n\tilde{\sigma}_Y R_f + n\tilde{\sigma}_X R_f - n\psi\tilde{\sigma}_X\tilde{\sigma}_Y + m\overline{y}\tilde{\sigma}_X}{(n+m)\tilde{\sigma}_X},
$$

$$\tag{6.8}$$

and Lagrangian

$$
\hat{\alpha} = \frac{\Sigma(y_j - \tilde{\mu}_Y)}{\tilde{\sigma}_Y}. \tag{6.9}
$$

So we have $\tilde{\theta} = (\tilde{\mu}_X, \tilde{\sigma}_X^2, \tilde{\mu}_Y, \tilde{\sigma}_Y^2)'$. We are now able to define the tilted log-likelihood:

$$
\tilde{l}(\theta) = l(\theta) + \hat{\alpha}\left[\frac{\mu_X - R_f}{\sigma_X} - \frac{\mu_Y - R_f}{\sigma_Y} - \psi\right], \tag{6.10}
$$

with first derivatives:

$$
\tilde{l}_{\mu_X} = l_{\mu_X} + \frac{\hat{\alpha}}{\sigma_X},
$$

$$
\tilde{l}_{\sigma_X^2} = l_{\sigma_X^2} - \hat{\alpha}\left[\frac{\mu_X - R_f}{2\sigma_X^3}\right],
$$

$$
\tilde{l}_{\mu_Y} = l_{\mu_Y} - \frac{\hat{\alpha}}{\sigma_Y},
$$

$$
\tilde{l}_{\sigma_Y^2} = l_{\sigma_Y^2} + \hat{\alpha}\left[\frac{\mu_Y - R_f}{2\sigma_Y^3}\right],
$$

$$
\tilde{l}_{\mu_X\mu_X} = l_{\mu_X\mu_X},
$$

$$
\tilde{l}_{\mu_X\sigma_X^2} = l_{\mu_X\sigma_X^2} - \frac{\hat{\alpha}}{2\sigma_Y^3},
$$

$$
\tilde{l}_{\mu_X\mu_Y} = 0.
$$

$$\tag{6.11}$$

The second derivatives are needed to calculate $\tilde{\jmath}_{\theta\theta}(\theta)$:

$$
\tilde{\jmath}_{\theta\theta}(\theta) = \jmath_{\theta\theta}(\theta) +
\begin{bmatrix}
0 & \dfrac{\widehat{\alpha}}{2\sigma_X^3} & 0 & 0 \\[2ex]
\dfrac{\widehat{\alpha}}{2\sigma_X^3} & -\dfrac{3\widehat{\alpha}(\mu_X - R_f)}{4\sigma_X^5} & 0 & 0 \\[2ex]
0 & 0 & 0 & -\dfrac{\widehat{\alpha}}{2\sigma_Y^3} \\[2ex]
0 & 0 & -\dfrac{\widehat{\alpha}}{2\sigma_Y^3} & \dfrac{3\widehat{\alpha}(\mu_Y - R_f)}{4\sigma_Y^5}
\end{bmatrix}.
\tag{6.12}
$$

The canonical parameter for this model is given by the following:

$$
\varphi_\theta(\theta) = \left( \frac{\mu_X}{\sigma_X^2}, \frac{1}{\sigma_X^2}, \frac{\mu_Y}{\sigma_Y^2}, \frac{1}{\sigma_Y^2} \right)'.
\tag{6.13}
$$

The matrices associated with this vector can now be calculated:

$$
\varphi_\theta(\theta) =
\begin{bmatrix}
\dfrac{1}{\sigma_X^2} & -\dfrac{\mu_X}{\sigma_X^4} & 0 & 0 \\[2ex]
0 & -\dfrac{1}{\sigma_X^4} & 0 & 0 \\[2ex]
0 & 0 & \dfrac{1}{\sigma_Y^2} & -\dfrac{\mu_Y}{\sigma_Y^4} \\[2ex]
0 & 0 & 0 & -\dfrac{1}{\sigma_Y^4}
\end{bmatrix},
\tag{6.14}
$$

$$
\varphi_\theta^{-1}(\theta) =
\begin{bmatrix}
\sigma_X^2 & -\mu_X\sigma_X^2 & 0 & 0 \\[1ex]
0 & -\sigma_X^4 & 0 & 0 \\[1ex]
0 & 0 & \sigma_Y^2 & -\mu_Y\sigma_Y^2 \\[1ex]
0 & 0 & 0 & -\sigma_Y^4
\end{bmatrix}.
$$

Given our parameter of interest

$$
\psi(\theta) = \frac{\mu_X - R_f}{\sqrt{\sigma_X^2}} - \frac{\mu_Y - R_f}{\sqrt{\sigma_Y^2}},
\tag{6.15}
$$

**Table 7:** 95% Confidence intervals for Sharpe ratio difference.

| Method | 95% CI |
|---|---|
| MLE | (−0.9857, 0.6953) |
| Likelihood ratio | (−0.9862, 0.6948) |
| Proposed | (−0.9739, 0.7057) |

we have

$$\psi_\theta(\theta) = \left( \frac{1}{\sqrt{\sigma_X^2}}, \frac{-(\mu_X - R_f)}{2(\sigma_X^2)^{3/2}}, \frac{-1}{\sqrt{\sigma_Y^2}}, \frac{(\mu_Y - R_f)}{2(\sigma_Y^2)^{3/2}} \right). \tag{6.16}$$

Quantities $\varphi_\theta(\theta)$, $\varphi_\theta^{-1}(\theta)$ and $\psi_\theta(\theta)$ can be used to calculate the new parameter $\chi(\theta)$ given in (3.18). The additional ingredients necessary for the construction of the standardized maximum likelihood departure $Q$ given in (3.19) are readily available from the above information.

## 7. Example and Simulations

In this section we provide an example and a simulation study for the two-sample Sharpe ratio case. We compare the third-order likelihood-based inference method to the classical methods used for testing, namely, the maximum likelihood statistic and the likelihood ratio statistic. These latter two statistics are analogous to expressions (3.4) and (3.5).

### 7.1. Example

For our example, we use the data presented in Table 1 to compare Sharpe ratios. We may for instance, be interested in whether the mutual fund's risk-adjusted return as captured by the Sharpe ratio is significantly better than the market's return. In Table 7 we present the 95% confidence interval for the difference between the Sharpe ratios for the mutual fund and market index. We can see that while the intervals produced using the MLE and likelihood ratio methods are similar, the interval obtained from the third-order method differs from these methods. As this is an example, we cannot comment on which interval is more accurate, but it is relevant to note the differences between the intervals which may be important in real world settings. The $p$ values for testing a null hypothesis of a zero difference between the Sharpe ratios of the mutual fund and market index are given by: 0.3675, 0.3675, and 0.3777 for the MLE, likelihood ratio and proposed, respectively. As tail probabilities tend to be small probabilities, it is important to approximate these as accurately as we can.

### 7.2. Simulations

In this section we provide a simulation study to assess the performance of the third-order method relative the MLE and likelihood ratio. The size of each simulation is 10,000 and again the parameter values were chosen to mimic the example data: $X \sim N(0.007378393, 0.01137605)$, $Y \sim N(0.01211267, 0.01789324)$, and $R_f = 0.003181247$. Confidence intervals for the difference between the Sharpe ratios for funds $X$ and $Y$ were obtained

**Table 8:** Results for Simulation Study.

| CI | $n, m$ | Method | Lower error | Upper error | Central coverage |
|---|---|---|---|---|---|
| | | MLE | 0.0616 | 0.0704 | 0.8680 |
| | | Likelihood ratio | 0.0619 | 0.0712 | 0.8669 |
| | $n = m = 12$ | Proposed | 0.0484 | 0.0539 | 0.8977 |
| | | *Nominal* | 0.0500 | 0.0500 | 0.9000 |
| | | *Standard error* | 0.0022 | 0.0022 | 0.0030 |
| | | MLE | 0.0645 | 0.0624 | 0.8731 |
| | | Likelihood ratio | 0.0651 | 0.0628 | 0.8721 |
| 90% | $n = 12, m = 20$ | Proposed | 0.0501 | 0.0498 | 0.9001 |
| | | *Nominal* | 0.0500 | 0.0500 | 0.9000 |
| | | *Standard error* | 0.0022 | 0.0022 | 0.0030 |
| | | MLE | 0.0550 | 0.0668 | 0.8782 |
| | | Likelihood ratio | 0.0550 | 0.0681 | 0.8769 |
| | $n = 20, m = 12$ | Proposed | 0.0508 | 0.0510 | 0.8982 |
| | | *Nominal* | 0.0500 | 0.0500 | 0.9000 |
| | | *Standard error* | 0.0022 | 0.0022 | 0.0030 |
| | | MLE | 0.0328 | 0.0381 | 0.9291 |
| | | Likelihood ratio | 0.0332 | 0.0393 | 0.9275 |
| | $n = m = 12$ | Proposed | 0.0235 | 0.0272 | 0.9493 |
| | | *Nominal* | 0.0250 | 0.0250 | 0.9500 |
| | | *Standard error* | 0.0016 | 0.0016 | 0.0022 |
| | | MLE | 0.0331 | 0.0340 | 0.9329 |
| | | Likelihood ratio | 0.0339 | 0.0345 | 0.9316 |
| 95% | $n = 12, m = 20$ | Proposed | 0.0254 | 0.0263 | 0.9483 |
| | | *Nominal* | 0.0250 | 0.0250 | 0.9500 |
| | | *Standard error* | 0.0016 | 0.0016 | 0.0022 |
| | | MLE | 0.0283 | 0.0360 | 0.9357 |
| | | Likelihood ratio | 0.0284 | 0.0381 | 0.9335 |
| | $n = 20, m = 12$ | Proposed | 0.0250 | 0.0259 | 0.9491 |
| | | *Nominal* | 0.0250 | 0.0250 | 0.9500 |
| | | *Standard error* | 0.0016 | 0.0016 | 0.0022 |
| | | MLE | 0.0085 | 0.0093 | 0.9822 |
| | | Likelihood ratio | 0.0094 | 0.0099 | 0.9807 |
| | $n = m = 12$ | Proposed | 0.0043 | 0.0061 | 0.9896 |
| | | *Nominal* | 0.0050 | 0.0050 | 0.9900 |
| | | *Standard error* | 0.0007 | 0.0007 | 0.0010 |
| | | MLE | 0.0070 | 0.0083 | 0.9847 |
| | | Likelihood ratio | 0.0081 | 0.0085 | 0.9834 |
| 99% | $n = 12, m = 20$ | Proposed | 0.0052 | 0.0046 | 0.9902 |
| | | *Nominal* | 0.0050 | 0.0050 | 0.9900 |
| | | *Standard error* | 0.0007 | 0.0007 | 0.0010 |
| | | MLE | 0.0066 | 0.0078 | 0.9856 |
| | | Likelihood ratio | 0.0066 | 0.0089 | 0.9845 |
| | $n = 20, m = 12$ | Proposed | 0.0053 | 0.0057 | 0.9890 |
| | | *Nominal* | 0.0050 | 0.0050 | 0.9900 |
| | | *Standard error* | 0.0007 | 0.0007 | 0.0010 |

for various sample sizes of each fund. Table 8 records the results from this simulation. Lower and upper probability errors are recorded as is central coverage. The nominal values and standard errors of the simulation are also reported for reference. As in the one sample case, these simulation results generally indicate that the propsed method outperforms the other methods based on the criteria we examined.

## 8. Conclusion

A higher order likelihood method was applied for inference on the Sharpe ratio. The two-sample problem for comparing Sharpe ratios was also considered. This statistical method is known to be extremely accurate and has $O(n^{-3/2})$ distributional accuracy. The methodology was demonstrated and worked out explicitly for independently and identically normally distributed returns. Simulations were provided to show the exceptional accuracy of the method even for very small sample sizes. While our assumption of independently and identically normally distributed returns may seem restrictive, we stress that the methodology can be applied more generally for any parametric return distribution structure and our choice of normality was made to clearly illustrate the methodology's merits to the applied practitioner. A natural next step would be to consider inference when returns are not independently and identically distributed. For the two-sample problem, it would be fruitful to compare the approach taken by Ledoit and Wolf [21] which involves the studentized bootstrap with the proposed methodology.

## Acknowledgment

## References

[1] W. F. Sharpe, "Mutual fund performance," *The Journal of Business*, vol. 39, no. 1, pp. 119–138, 1966.

[2] C. Keating and W. F. Shadwick, *An Introduction to Omega*, The Finance Development Centre, 2002.

[3] M. Eling and F. Schuhmacher, "Does the choice of performance measure influence the evaluation of hedge funds?" *Journal of Banking and Finance*, vol. 31, no. 9, pp. 2632–2647, 2007.

[4] J. D. Opdyke, "Comparing Sharpe ratios: so where are the p-values?" *Journal of Asset Management*, vol. 8, no. 5, pp. 308–336, 2007.

[5] J. D. Jobson and B. M. Korkie, "Performance hypothesis testing with the Sharpe and Treynor measures," *The Journal of Finance*, vol. 36, no. 4, pp. 889–908, 1981.

[6] A. W. Lo, "The statistics of Sharpe ratios," *Financial Analysts Journal*, vol. 58, no. 4, pp. 36–52, 2002.

[7] D. R. Cox and D. V. Hinkley, *Theoretical Statistics*, Chapman and Hall, London, UK, 1974.

[8] E. Mertens, "Comments on variance of the IID estimator in Lo," Research Note, 2002, http://www.elmarmertens.org/.

[9] S. Christie, "Is the Sharpe ratio useful in asset allocation?" MAFC Research Papers 31, Applied Finance Centre, Macquarie University, 2005.

[10] N. Doganaksoy and J. Schmee, "Comparisons of approximate confidence intervals for distributions used in life-data analysis," *Technometrics*, vol. 35, no. 2, pp. 175–184, 1993.

[11] H. E. Daniels, "Saddlepoint approximations in statistics," *Annals of Mathematical Statistics*, vol. 25, pp. 631–650, 1954.

[12] H. E. Daniels, "Tail probability approximations," *International Statistical Review*, vol. 55, no. 1, pp. 37–48, 1987.

[13] O. E. Barndorff-Nielsen and D. R. Cox, *Inference and Asymptotics*, vol. 52 of *Monographs on Statistics and Applied Probability*, Chapman & Hall, London, UK, 1994.

[14] O. E. Barndorff-Nielsen, "Conditionality resolutions," *Biometrika*, vol. 67, no. 2, pp. 293–310, 1980.

[15] D. A. S. Fraser, N. Reid, and A. Wong, "Exponential linear models: a two-pass procedure for saddle-point approximation," *Journal of the Royal Statistical Society B*, vol. 53, no. 2, pp. 483–492, 1991.

[16] R. Lugannani and S. Rice, "Saddlepoint approximation for the distribution of the sum of independent random variables," *Advances in Applied Probability*, vol. 12, no. 2, pp. 475–490, 1980.

[17] O. E. Barndorff-Nielsen, "Inference on full or partial parameters based on the standardized signed log likelihood ratio," *Biometrika*, vol. 73, no. 2, pp. 307–322, 1986.

[18] J. L. Jensen, "The modified signed likelihood statistic and saddlepoint approximations," *Biometrika*, vol. 79, no. 4, pp. 693–703, 1992.

[19] D. A. S. Fraser and N. Reid, "Ancillaries and third order significance," *Utilitas Mathematica*, vol. 47, pp. 33–53, 1995.

[20] D. A. S. Fraser, N. Reid, and J. Wu, "A simple general formula for tail probabilities for frequentist and Bayesian inference," *Biometrika*, vol. 86, no. 2, pp. 249–264, 1999.

[21] O. Ledoit and M. Wolf, "Robust performance hypothesis testing with the Sharpe ratio," *Journal of Empirical Finance*, vol. 15, no. 5, pp. 850–859, 2008.