entropy

MDPI

Article

Adjusted Empirical Likelihood Method in the Presence of Nuisance Parameters with Application to the Sharpe Ratio

Yuejiao Fu, Hangjing Wang and Augustine Wong *

Department of Mathematics and Statistics, York University, 4700 Keele Street, Toronto, ON M3J 1P3, Canada; yuejiao@mathstat.yorku.ca (Y.F.); hangjing@mathstat.yorku.ca (H.W.)

* Correspondence: august@yorku.ca

Received: 24 February 2018; Accepted: 11 April 2018; Published: 25 April 2018

Abstract: The Sharpe ratio is a widely used risk-adjusted performance measurement in economics and finance. Most of the known statistical inferential methods devoted to the Sharpe ratio are based on the assumption that the data are normally distributed. In this article, without making any distributional assumption on the data, we develop the adjusted empirical likelihood method to obtain inference for a parameter of interest in the presence of nuisance parameters. We show that the log adjusted empirical likelihood ratio statistic is asymptotically distributed as the chi-square distribution. The proposed method is applied to obtain inference for the Sharpe ratio. Simulation results illustrate that the proposed method is comparable to Jobson and Korkie's method (1981) and outperforms the empirical likelihood method when the data are from a symmetric distribution. In addition, when the data are from a skewed distribution, the proposed method significantly outperforms all other existing methods. A real-data example is analyzed to exemplify the application of the proposed method.

Keywords: adjusted empirical likelihood; coverage probability; nonparametric; nuisance parameter; Sharpe ratio

1. Introduction

In financial economics, Sharpe ratio, defined in [1], provides a measure of a fund's excess returns relative to its volatility. Let μ be an expected return of an asset, and σ be the corresponding standard deviation. The Sharpe ratio is defined as

$$sr = \frac{\mu - R_f}{\sigma},$$

where R_f is a known risk-free rate of return. Note that the larger the Sharpe ratio is, the more return the investor is getting per unit of risk. It is the standard convention in economics and finance research to report the Sharpe ratio. Therefore, the Sharpe ratio is very well studied as a measure of the mutual fund performance in the financial economic areas such as the portfolio analysis, the pricing of capital asset under conditions of risk and the general behavior of stock market prices. The popularity of the Sharpe ratio in financial economics is not only from its simplicity; the study of the Sharpe ratio will also directly result in deeper understandings in portfolio selections. Assuming that the asset returns are all normally distributed, Sharpe [1] showed that picking an asset with the largest Sharpe ratio is equivalent to finding a solution of the investor's expected utility problem.

Under the normality assumption, Jobson and Korkie [2] proposed a parametric test for the Sharpe ratio, which is a very popular inferential method in economics and finance. However, as shown by many researchers [3–6], it is very common for the actual returns of the investments, such as the hedge

funds, to have a skewed distribution. When the normality assumption of the investment returns is violated, the commonly used approximate distributions of the Sharpe ratio which are developed under the normality assumption become problematic. Model mis-specification is a big concern for all parametric approaches since a misspecified model may lead to biased results. Since the Sharpe ratio is only involved in the first two moments of the data, one of the themes attempting to resolve the problem is to consider higher order moments. There was abundant literature along this line of research such as [7–10] and references therein.

Another line of research to the problem is to use the nonparametric approach. In this article, we adopt the empirical likelihood (EL) method. Empirical likelihood-type method was first used by Thomas and Grunkemeier [11] to study the survival probabilities estimated by the Kaplan–Meier curve. Owen [12,13] formalized the EL as a unified inference method under more general settings. The EL-based confidence region has several beneficial properties: it does not impose prior constraints on region shape, is transformation invariant and Bartlett correctable [14]. Qin and Lawless [15] applied the EL to inference on parameters that are generated from estimating equations. When the sample size is small and/or the dimension of the estimating equations is high, the EL approach can be hindered by an empty set problem and under-coverage problem. In order to resolve the empty set problem and improve the coverage probability of the statistical tests of the ordinary EL methods, Chen et al. [16] proposed the adjusted empirical likelihood (AEL) method by adding one artificial point into the data set. However, only problems without nuisance parameters were considered in [16]. In this article, we focus on the AEL method with nuisance parameters in addition to the parameter of primary interest. We develop the asymptotic theory of the AEL method when nuisance parameters exist, and demonstrate the use of the AEL method in the application of the Sharpe ratio. Our simulation studies show that the proposed approach provides a beneficial robust alternative to the inference of the Sharpe ratio. The proposed AEL method is comparable to Jobson and Korkie's method [2] and outperforms the EL method when the data are from a symmetric distribution, while for data generating from a skewed distribution, the proposed method outperforms all other existing methods, especially for small sample sizes. The AEL method preserves the advantage of the EL method: the shape of confidence region based on the AEL ratio reflects the observed data set, while the confidence region based on other methods (excluding EL) is always symmetric about the point estimator. Therefore, the AEL approach allows the data to speak for themselves, and is robust against model mis-specification.

The rest of the article is organized as follows. A brief introduction to the EL and AEL methodologies is given in Section 2. In Section 3, we study the asymptotic property of the AEL method with nuisance parameters. In Section 4, simulation studies are conducted to investigate the precision of the coverage probabilities in the context of the Sharpe ratio. In Section 5, a real-data example is analyzed to illustrate the application of the proposed method. Some concluding remarks are given in Section 6. The technical details are presented in the Appendix.

2. Review of the Empirical Likelihood and the Adjusted Empirical Likelihood Methods

Let $X_1, X_2, \ldots, X_n \in \mathbb{R}^d$ be the independent and identically distributed random vectors following distribution F with mean μ and a nonsingular covariance matrix. The corresponding observed values are denoted by $x_1, x_2, \ldots, x_n$. The EL function for the population distribution F is given by

$$L(F) = \prod_{i=1}^{n} F(\{x_i\}),$$

where $F(\{x_i\})$ is the probability of observing the value x_i in a sample from F. Denote $p_i = F(\{x_i\})$. The EL function can also be written as

$$L(F) = \prod_{i=1}^{n} p_i. \tag{1}$$

Clearly , we have $0 \le p_i \le 1$ and $\sum_{i=1}^n p_i = 1$. Suppose that the goal is to construct a confidence region for the mean μ. The profile EL function of μ is defined to be

$$L_{EL}(\mu) = \sup\Big\{\prod_{i=1}^n p_i : \ p_i \ge 0, i = 1, \ldots, n; \ \sum_{i=1}^n p_i = 1; \ \sum_{i=1}^n p_i x_i = \mu\Big\}.$$

Qin and Lawless [15] showed that extra information in the form of a set of estimating equations can be used to improve the maximum empirical likelihood estimators (MELE) and the EL ratio confidence intervals. Suppose a k dimensional parameter θ is associated with F via a vector $g(x, \theta)$ of $r \ge k$ functionally independent unbiased estimating functions. Then for each $j = 1, 2, \ldots, r$, we have an estimating equation $E_F\{g_j(x, \theta)\} = 0$, which can be written in the vector form as $E_F\{g(x, \theta)\} = 0$. The profile EL function of θ is

$$L_{EL}(\theta) = \sup\Big\{\prod_{i=1}^n p_i : \ p_i \ge 0, i = 1, \ldots, n; \ \sum_{i=1}^n p_i = 1; \ \sum_{i=1}^n p_i g(x_i, \theta) = 0\Big\}, \tag{2}$$

and hence, the profile log-EL function is

$$l_{EL}(\theta) = \sup\Big\{\sum_{i=1}^n \log p_i : \ p_i \ge 0, i = 1, \ldots, n; \ \sum_{i=1}^n p_i = 1; \ \sum_{i=1}^n p_i g(x_i, \theta) = 0\Big\}. \tag{3}$$

The constrained optimization problem in (3) can be solved by applying the method of Lagrange multipliers. Let λ and $t = (t_1, \ldots, t_r)^\tau$ be Lagrange multipliers and define

$$H = \sum_i \log p_i + \lambda(1 - \sum_i p_i) - nt^\tau \sum_i p_i g(x_i, \theta). \tag{4}$$

Then maximizing (3) is equivalent to maximizing H unconditionally. Setting the first partial derivative of (4) with respect to p_i equal to 0, we have

$$\frac{\partial H}{\partial p_i} = \frac{1}{p_i} - \lambda - nt^\tau g(x_i, \theta) = 0,$$

$$\sum_{i=1}^n p_i \frac{\partial H}{\partial p_i} = n - \lambda = 0 \quad \Rightarrow \quad \lambda = n$$

and

$$\hat{p}_i = \frac{1}{n[1 + t^\tau g(x_i, \theta)]},$$

where t can be expressed as a function of θ by solving the following equations

$$\sum_{i=1}^n \hat{p}_i g(x_i, \theta) = 0. \tag{5}$$

Now the profile log-EL function can be written as

$$l_{EL}(\theta) = -\sum_{i=1}^n \log\Big[1 + t^\tau g(x_i, \theta)\Big] - n \log n. \tag{6}$$

Note that (5) can be rewritten as

$$\sum_{i=1}^n \frac{g(x_i, \theta)}{1 + t^\tau g(x_i, \theta)} = 0. \tag{7}$$

Now maximizing (3) has been transformed into an equivalence of solving (7) for the Lagrange multiplier t. In practice, this is achieved by numerical methods. One such algorithm devoted to this end can be found in [16]. A necessary and sufficient condition for the existence of a solution $\tilde{t} = \tilde{t}(\theta)$ in (7) is that 0 must be an inner point of the convex hull expanded by $\{g(x_i, \theta), i = 1, 2, \ldots, n\}$.

Qin and Lawless [15] further showed that under some regularity conditions, the EL ratio statistic $W_0(\theta_0) = 2[l_{EL}(\tilde{\theta}) - l_{EL}(\theta_0)]$ converges to χ^2_k in distribution as the sample size n approaches infinity. This result is the foundation for hypothesis test on θ and can be used to construct an approximate $100(1-\alpha)\%$ confidence region of θ,

$$I_{EL} = \{\mu : W_0(\theta) \leq \chi^2_k(1-\alpha)\},$$

where $\chi^2_k(1-\alpha)$ is the $100(1-\alpha)\%$ quantile of the χ^2_k distribution, and α is a pre-specified significance level.

Under mild conditions, the convex hull of $\{g(x_i, \theta),\ i = 1, 2, \ldots, n\}$ contains 0 as its inner point with probability 1 as $n \to \infty$. However, if θ is not close to the true parameter θ_0 or when the sample size n is small, the convex hull is not guaranteed to contain 0. Thus, there is a nonzero probability that the solution to (7) does not exist. It results computational issues when solving the constrained optimization problem in the definition of the EL function. This is known as the empty set problem or the convex hull problem in the EL literature.

In order to resolve the convex hull problem, Chen et al. [16] proposed the AEL method by adding one artificial point into the data set. Denote

$$g_i = g_i(\theta) = g(x_i, \theta)$$

and

$$\bar{g}_n = \bar{g}_n(\theta) = \frac{1}{n}\sum_{i=1}^{n} g_i.$$

Let $a_n = o(n)$ be a given positive constant. Define a new point by

$$g_{n+1} = g_{n+1}(\theta) = -\frac{a_n}{n}\sum_{i=1}^{n} g_i = -a_n\bar{g}_n.$$

Similar to (2), the profile log-AEL function if defined as

$$l_{AEL}(\theta) = \sup\Big\{\sum_{i=1}^{n+1}\log[(n+1)p_i] :\ p_i \geq 0, i = 1, \ldots, n+1;\ \sum_{i=1}^{n+1} p_i = 1;\ \sum_{i=1}^{n+1} p_i g_i = 0\Big\},$$

and we have

$$l_{AEL}(\theta) = -\sum_{i=1}^{n+1}\log\Big[1 + t^\tau g(x_i, \theta)\Big], \tag{8}$$

where t satisfies

$$\sum_{i=1}^{n+1}\frac{g(x_i, \theta)}{1 + t^\tau g(x_i, \theta)} = 0. \tag{9}$$

The introduction of g_{n+1} guarantees a solution for t in (7). Let the maximum AEL estimator $\tilde{\theta}$ be the maximizer of $l_{AEL}(\theta)$. Under mild regularity conditions, the AEL ratio statistic $W(\theta_0) = 2[l_{AEL}(\tilde{\theta}) - l_{AEL}(\theta_0)]$ converges to χ^2_k in distribution as the sample size n approaches infinity. Chen et al. [16] showed that the statistical tests based on the AEL method give better coverage probabilities than those obtained by the original EL method.

In this article, we propose using the AEL method to conduct inference on the Sharpe ratio. Suppose the data is from a population with mean μ and variance σ^2. Without loss of generality, for the rest of this article, define the Sharpe ratio of the population as

$$sr = \frac{\mu}{\sigma}.$$

In this case, the parameter vector is $\theta = (\mu, \sigma^2)$, and the parameter of interest is sr. The set of estimating functions can either be

$$X - \mu \quad \text{and} \quad (X-\mu)^2 - \left(\frac{\mu}{sr}\right)^2 \tag{10}$$

or

$$X - \sigma(sr) \quad \text{and} \quad (X - \sigma(sr))^2 - \sigma^2, \tag{11}$$

which has μ or σ as the nuisance parameter, respectively. Chen et al. [16] discussed the AEL-based inference without nuisance parameters. Building upon [15,16], we develop the convergence theorem for the AEL with nuisance parameters as shown in the next section.

3. The Adjusted Empirical Likelihood Method in the Presence of Nuisance Parameters

Suppose a k dimensional parameter $\theta = (\theta_1, \theta_2)$ consists a q dimensional parameter of interest θ_1 as well as a $(k-q)$ dimensional nuisance parameter θ_2. The goal is to test $H_0 : \theta_1 = \theta_1^0$ for some given θ_1^0. In order to obtain inference for θ_1 using the AEL method, the asymptotic results in [16] need to be reconstructed and extended to the situation with nuisance parameters.

First, we develop a lemma about positive definite matrices. If a matrix M is positive semidefinite, we denote it by $M \geq 0$; if M is positive definite, we write $M > 0$. For any matrices G and H, let $G \geq H$ denote that $G - H$ is positive semidefinite, and let $G > H$ denote that $G - H$ is positive definite.

Lemma 1. *Let M be a $k \times k$ symmetric positive definite block matrix of the form*

$$M = \begin{pmatrix} A & B \\ B^\tau & C \end{pmatrix},$$

where A is a $q \times q$ matrix, B is a $q \times (k-q)$ matrix, and C is a $(k-q) \times (k-q)$ matrix. Then C is positive definite and

$$\begin{pmatrix} A & B \\ B^\tau & C \end{pmatrix}^{-1} \geq \begin{pmatrix} 0 & 0 \\ 0 & C^{-1} \end{pmatrix}.$$

The proof of the above lemma is given in Appendix. In order to prove the main theorem, we also need the following two results about idempotent matrices. The proof of these two results can be found in [17] (pp. 186–187).

Result 1. *A necessary and sufficient condition that $Y'AY$ has a χ^2 distribution is that A is idempotent, that is, $A^2 = A$, in which case the degrees of freedom of χ^2 is rank A = trace A.*

Result 2. *If A, B, $A - B$ are matrices of non-negative quadratic forms and A and B are idempotent, then $A - B$ is also idempotent.*

Based on Lemma (1) and the above two results, we have the following theorem which gives the asymptotic properties of the AEL ratio test statistic. The theorem is a nonparametric analogue of the theorem in [18] on the asymptotic distribution of the likelihood ratio. The difference is that Wilks' theorem is based on parametric likelihood and ours is based on the adjusted empirical likelihood.

Moreover, it takes into consideration nuisance parameters. We follow the idea of profiling out nuisance parameters (Corollary 5 in [15] and Corollary 1 in [19]) to perform the AEL ratio test. The proof of the theorem is provided in Appendix.

Theorem 1. *Let $\theta^\tau = (\theta_1, \theta_2)^\tau$, where θ_1 and θ_2 are $q \times 1$ and $(k-q) \times 1$ vectors, respectively. For $H_0 : \theta_1 = \theta_1^0$, the profile AEL ratio test statistic is*

$$W(\theta_1^0) = 2[l_{AEL}(\tilde{\theta}_1, \tilde{\theta}_2) - l_{AEL}(\theta_1^0, \tilde{\theta}_2^0)],$$

where $\tilde{\theta}^\tau = (\tilde{\theta}_1, \tilde{\theta}_2)^\tau$ maximizes $l_{AEL}(\theta) = l_{AEL}(\theta_1, \theta_2)$, and $\tilde{\theta}_2^0$ maximizes $l_{AEL}(\theta_1^0, \theta_2)$ with respect to θ_2. Under H_0, $W(\theta_1^0) \xrightarrow{d} \chi_q^2$ as $n \to \infty$.

It is worth noticing that Theorem 1 holds true as long as $a_n = o_p(n)$. In application, a_n with higher orders is usually not recommended, since the AEL ratios are decreasing functions of the adjustment level a_n [20]. As suggested by [16], we set $a_n = \frac{1}{2} \log n$ for all of the simulations and applications if not otherwise specified.

Since in Theorem 1 , θ_1 is the parameter of interest and θ_2 is considered as the nuisance parameter. We can apply the theorem to the Sharpe ratio by setting $\theta_1 = sr$ along with $\theta_2 = \mu$ or $\theta_2 = \sigma^2$. Therefore, the AEL ratio statistic under the null hypothesis $H_0 : sr = sr_0$ can be either

$$W(sr_0) = 2[l_{AEL}(\tilde{sr}, \tilde{\mu}) - l_{AEL}(sr_0, \tilde{\mu}_0)], \tag{12}$$

or

$$W(sr_0) = 2[l_{AEL}(\tilde{sr}, \tilde{\sigma}^2) - l_{AEL}(sr_0, \tilde{\sigma}_0^2)]. \tag{13}$$

Our simulation shows that using (12) or (13) as the AEL ratio statistic does not make any significant difference in the inference of sr.

4. Simulation Study

In order to evaluate the accuracy of the asymptotic chi-square calibration of the AEL method, we choose the coverage probability as an indicator throughout this section. For some fixed sample size n and sr_0, suppose we have run the simulation m times and s of the simulated $W(sr_0)$ are less than the $1-\alpha$ quantile of χ_1^2 for some given α. Then the coverage probability is defined to be s/m, which is compared with the nominal value $1-\alpha$. When m is large, if the coverage probability s/m is close to $1-\alpha$, then the level α test for sr will tend to give good performance and χ_1^2 is considered an acceptable reference distribution for $W(sr_0)$ at sample size n.

We compare the coverage probability of the proposed method with other methods for sample sizes $n = 20, 50, 200, 500$ at nominal values $1-\alpha = 0.9, 0.95$. Each coverage probability is obtained from $m = 5000$ simulations. The data are generated from the normal distribution with mean $\mu = 1$ and standard deviation $\sigma = 0.5$, t-distribution and the chi-square distributions with various degrees of freedom. The methods under comparison are the following: the Jobson and Korkie's method [2] (JK), the Mertens's method [21] (Mertens), the usual EL inferential method (EL), application of the delta method on the asymptotic distribution of the EL estimator of the mean and standard deviation (Delta), and the proposed method (AEL) with the adjustment level $a_n = 0.5 \log n$. Jobson and Korkie [2] assumed that the data are from a normal distribution. By applying the delta method to approximate the mean and variance of the Sharpe ratio, confidence interval for the Sharpe ratio can then be approximated by the Central Limit Theorem. Mertens [21] used the skewness and kurtosis to give an adjusted approximation of the variance of the Sharpe ratio derived in Jobson and Korkie [2] and again obtained the confidence interval of the Sharpe ratio from the Central Limit Theorem. The approach denoted by Delta is similar to JK but based on the EL. For the EL method, whenever the convex hull problem occurs for a set of

simulated data, we use the convention to set the value of the profile log-EL function as negative infinity. Results are summarized in Table 1.

Table 1. Coverage probabilities of the Sharpe ratio.

$1-\alpha$	Method	$n=20$	$n=50$	$n=200$	$n=500$
		$N(1, 0.25)$			
0.9	JK	0.8956	0.9022	0.8968	0.9060
	Mertens	0.8258	0.8694	0.8894	0.9004
	EL	0.8210	0.8760	0.8906	0.9040
	AEL	0.8486	0.8874	0.8942	0.9058
	Delta	0.8428	0.8840	0.8896	0.8976
0.95	JK	0.9460	0.9488	0.9488	0.9522
	Mertens	0.8926	0.9270	0.9414	0.9494
	EL	0.8762	0.9214	0.9440	0.9514
	AEL	0.8980	0.9312	0.9466	0.9534
	Delta	0.9054	0.9334	0.9408	0.9476
		t_3			
0.9	JK	0.8960	0.9004	0.9030	0.9040
	Mertens	0.8390	0.8646	0.8782	0.8890
	EL	0.8428	0.8738	0.8884	0.8946
	AEL	0.8794	0.8896	0.8944	0.8976
	Delta	0.8240	0.8586	0.8766	0.8858
0.95	JK	0.9494	0.9538	0.9516	0.9550
	Mertens	0.9028	0.9144	0.9326	0.9442
	EL	0.9042	0.9268	0.9438	0.9508
	AEL	0.9340	0.9402	0.9466	0.9514
	Delta	0.8910	0.9092	0.9318	0.9372
		t_6			
0.9	JK	0.8982	0.8984	0.8934	0.8976
	Mertens	0.8738	0.8840	0.8900	0.8954
	EL	0.8700	0.8860	0.8928	0.8962
	AEL	0.8986	0.9018	0.8966	0.8978
	Delta	0.8634	0.8828	0.8936	0.9008
0.95	JK	0.9504	0.9482	0.9466	0.9476
	Mertens	0.9246	0.9364	0.9426	0.9458
	EL	0.9240	0.9394	0.9438	0.9466
	AEL	0.9466	0.9470	0.9480	0.9481
	Delta	0.9214	0.9330	0.9460	0.9494
		χ^2_4			
0.9	JK	0.9640	0.9532	0.9476	0.9474
	Mertens	0.8048	0.8474	0.8676	0.8938
	EL	0.7800	0.8352	0.8626	0.8942
	AEL	0.8216	0.8536	0.8664	0.8952
	Delta	0.8072	0.8354	0.8660	0.8914
0.95	JK	0.9872	0.9808	0.9808	0.9774
	Mertens	0.8780	0.9072	0.9278	0.9422
	EL	0.8562	0.8972	0.9194	0.9418
	AEL	0.8924	0.9126	0.9228	0.9430
	Delta	0.8872	0.9046	0.9252	0.9388
		χ^2_6			
0.9	JK	0.9466	0.9476	0.9392	0.9414
	Mertens	0.8048	0.8460	0.8760	0.8916
	EL	0.7996	0.8392	0.8728	0.8904
	AEL	0.8346	0.8568	0.8780	0.8926
	Delta	0.8236	0.8524	0.8766	0.8846
0.95	JK	0.9796	0.9776	0.9758	0.9754
	Mertens	0.8780	0.9168	0.9352	0.9412
	EL	0.8626	0.9032	0.9284	0.9402
	AEL	0.8894	0.9156	0.9326	0.9410
	Delta	0.8916	0.9170	0.9336	0.9432

From Table 1, we can see that the AEL method has the most robust performance for various underlying population distributions. The AEL method always has significantly better performance over the EL method in terms of coverage probability. When the data is normally distributed, the JK method performs the best while when the data comes from a skewed distribution, the JK method performs poorly. For normal data with small sample size, the AEL has slightly less coverage probabilities than the JK method, while for normal data with sample size larger than 50 and data from various t distributions, the AEL has comparable performance with the JK method. For all other situations, the AEL method significantly outperforms all other methods, especially for cases with small sample sizes.

5. Real Data Analysis

The data we consider is the Nasdaq GS return of the Apple Inc. (Cupertino, CA, USA) from 3 October 2017 to 12 December 2017 (https://finance.yahoo.com/quote/AAPL/). The return is evaluated from the close price of the current day compared with the close price of the previous day. There are 50 trading days during the period considered. We use the yearly return rate of the 5-year bonds, which is 2.116%, as the yearly risk-free return. Therefore, the daily risk-free return rate used in the analysis is $0.02116/252 = 8.397 \times 10^{-5}$. Based on our data, the Durbin-Watson test statistic is 1.58. Hence, there is no significant evidence of serial correlation. The qqplot of the returns in Figure 1 reveals some skewness of the data. The confidence intervals of the Sharpe ratio for the Apple Inc. return data produced by different methods are listed in Table 2. For JK and Mertens methods, the point estimates are the value of sr that corresponding to the 50% quantile of the standard normal limiting distribution of their test statistics. The estimates of the Delta, EL and AEL methods are the value of the maximum EL and AEL estimates, respectively.

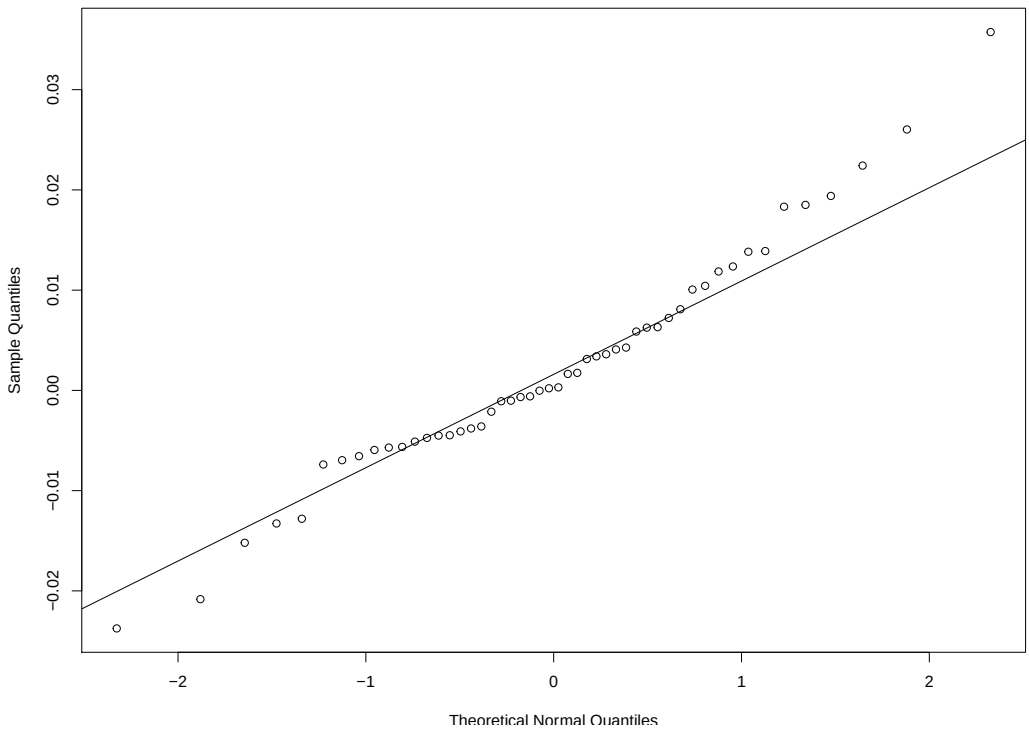

Figure 1. Quantile-quantile plot of Apple Inc. return data.

From Table 2, we see that since JK and Mertens methods are moment-based methods, both their estimates are the same as the sample Sharpe ratio. The Delta, EL and AEL methods are empirical-likelihood-based methods so the corresponding estimates are different from the previous two approaches. We observe that there is some difference in the confidence intervals for various

approaches. Note that the data has some skewness as shown in Figure 1. Based on the observation from our simulation studies, the skewness will affect the JK method but not the rest of the four methods. The confidence interval based on our proposed AEL method is more robust and trustworthy.

Table 2. Confidence Intervals of the Sharpe ratio for Apple Inc. return data.

$1-\alpha$	Method	Estimate	Lower Bound	Upper Bound
0.9	JK	0.1907	−0.0441	0.4254
	Mertens	0.1907	−0.0350	0.4163
	Delta	0.1926	−0.0329	0.4181
	EL	0.1926	−0.0376	0.4140
	AEL	0.1926	−0.0479	0.4241
0.95	JK	0.1907	−0.0890	0.4703
	Mertens	0.1907	−0.0783	0.4596
	Delta	0.1926	−0.0761	0.4613
	EL	0.1926	−0.0827	0.4558
	AEL	0.1926	−0.0949	0.4683

6. Conclusions

We extended the adjusted empirical likelihood method [16] to obtain inference for the a parameter of interest in the presence of nuisance parameters. The advantage of the proposed method is that it does not rely on the distributional assumption of the data. In particular, we applied the proposed method to obtain inference for the Sharpe ratio. Simulation results show that the proposed method gives the coverage probabilities closest to the nominal value than those obtained by the standard empirical likelihood ratio method. Simulation results illustrate that the proposed method is comparable to Jobson and Korkie's method [2] and outperforms the EL method when the data are from a symmetric distribution. In addition, when the data are from a skewed distribution, the proposed method outperforms all other existing methods.

The time-series properties of investment strategies can have a nontrivial impact on the Sharpe ratio estimator. In this article, we proposed using empirical-likelihood-based inference for Sharpe ratio. Empirical likelihood was motivated by independent and identically distributed data. When dealing with dependent data, we need to account for the dependency structure in constructing confidence regions for the parameter of interest. In general, the approach to handle dependent data within the EL framework is parallel to the methods based on parametric likelihood. The extension of our approach for dependent data is valuable and interesting. We will consider it in future research.

Author Contributions: All authors equally participated in the design, methodology, writing and interpretation of the results. H.W. conducted the analysis in R as part of her Ph.D. thesis. All authors have read and approved the final manuscript.

Acknowledgments: The authors would like to thank the Editor and three referees for their valuable suggestions and comments.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix

Proof of Lemma 1. Since M is a symmetric positive matrix, we have $C > 0$ and $A - BC^{-1}B^{\tau} > 0$; see Theorem 16.1 in [22]. Noting that M has the following factorization

$$\begin{pmatrix} A & B \\ B^{\tau} & C \end{pmatrix} = \begin{pmatrix} I & BC^{-1} \\ 0 & I \end{pmatrix} \begin{pmatrix} A - BC^{-1}B^{\tau} & 0 \\ 0 & C \end{pmatrix} \begin{pmatrix} I & 0 \\ (BC^{-1})^{\tau} & I \end{pmatrix},$$

we have

$$\begin{pmatrix} A & B \\ B^\tau & C \end{pmatrix}^{-1} = \begin{pmatrix} I & 0 \\ (BC^{-1})^\tau & I \end{pmatrix}^{-1} \begin{pmatrix} (A - BC^{-1}B^\tau)^{-1} & 0 \\ 0 & C^{-1} \end{pmatrix} \begin{pmatrix} I & BC^{-1} \\ 0 & I \end{pmatrix}^{-1}.$$

Further note that

$$\begin{aligned} &\begin{pmatrix} 0 & 0 \\ 0 & C^{-1} \end{pmatrix} \\ =&\begin{pmatrix} I & 0 \\ (BC^{-1})^\tau & I \end{pmatrix}^{-1} \begin{pmatrix} I & 0 \\ (BC^{-1})^\tau & I \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & C^{-1} \end{pmatrix} \begin{pmatrix} I & BC^{-1} \\ 0 & I \end{pmatrix} \begin{pmatrix} I & BC^{-1} \\ 0 & I \end{pmatrix}^{-1} \\ =&\begin{pmatrix} I & 0 \\ (BC^{-1})^\tau & I \end{pmatrix}^{-1} \begin{pmatrix} 0 & 0 \\ 0 & C^{-1} \end{pmatrix} \begin{pmatrix} I & BC^{-1} \\ 0 & I \end{pmatrix}^{-1}. \end{aligned}$$

Above two factorizations lead to

$$\begin{aligned} &\begin{pmatrix} A & B \\ B^\tau & C \end{pmatrix}^{-1} - \begin{pmatrix} 0 & 0 \\ 0 & C^{-1} \end{pmatrix} \\ =&\begin{pmatrix} I & 0 \\ (BC^{-1})^\tau & I \end{pmatrix}^{-1} \left[\begin{pmatrix} (A - BC^{-1}B^\tau)^{-1} & 0 \\ 0 & C^{-1} \end{pmatrix} - \begin{pmatrix} 0 & 0 \\ 0 & C^{-1} \end{pmatrix} \right] \begin{pmatrix} I & BC^{-1} \\ 0 & I \end{pmatrix}^{-1} \\ =&\begin{pmatrix} I & 0 \\ (BC^{-1})^\tau & I \end{pmatrix}^{-1} \begin{pmatrix} (A - BC^{-1}B^\tau)^{-1} & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} I & BC^{-1} \\ 0 & I \end{pmatrix}^{-1}. \end{aligned}$$

Since $A - BC^{-1}B^\tau > 0$, we have $(A - BC^{-1}B^\tau)^{-1} > 0$, which leads to

$$\begin{pmatrix} A & B \\ B^\tau & C \end{pmatrix}^{-1} \geq \begin{pmatrix} 0 & 0 \\ 0 & C^{-1} \end{pmatrix}.$$

□

Proof of Theorem 1. For simplicity, denote $l(\theta) = -l_{AEL}(\theta)$. Then $\tilde{\theta}^\tau = (\tilde{\theta}_1, \tilde{\theta}_2)^\tau$ minimizes $l(\theta) = l(\theta_1, \theta_2)$, and $\tilde{\theta}_2^0$ minimizes $l(\theta_1^0, \theta_2)$ with respect to θ_2. Under this new notation, the test statistic becomes

$$W(\theta_1^0) = 2[l(\theta_1^0, \tilde{\theta}_2^0) - l(\tilde{\theta}_1, \tilde{\theta}_2)].$$

First, the following notations are needed in this proof. Let

$$Q_{1n}(\theta, t) = \frac{1}{n+1} \sum_{i=1}^{n+1} \frac{g_i(\theta)}{1 + t^\tau g_i(\theta)},$$

$$Q_{2n}(\theta, t) = \frac{1}{n+1} \sum_{i=1}^{n+1} \frac{1}{1 + t^\tau g_i(\theta)} \left(\frac{\partial g_i(\theta)}{\partial \theta} \right)^\tau t.$$

Let $\tilde{\theta}$ and $\tilde{t} = t(\tilde{\theta})$ be the solution of

$$Q_{1n}(\tilde{\theta}, \tilde{t}) = 0, \quad Q_{2n}(\tilde{\theta}, \tilde{t}) = 0.$$

The existence of $\tilde{\theta}$ and $\tilde{t} = t(\tilde{\theta})$ in a neighborhood of the true parameter θ_0 is proved in [15,16]. Note that

$$\frac{\partial Q_{1n}(\theta,0)}{\partial\theta} = \frac{1}{n+1}\sum_{i=1}^{n+1}\frac{\partial g_i(\theta)}{\partial\theta}, \qquad \frac{\partial Q_{1n}(\theta,0)}{\partial t^\tau} = -\frac{1}{n+1}\sum_{i=1}^{n+1} g_i(\theta)g_i(\theta)^\tau,$$

$$\frac{\partial Q_{2n}(\theta,0)}{\partial\theta} = 0, \qquad \frac{\partial Q_{2n}(\theta,0)}{\partial t^\tau} = \frac{1}{n+1}\sum_{i=1}^{n+1}\Big(\frac{\partial g_i(\theta)}{\partial\theta}\Big)^\tau.$$

Taylor expansion of $Q_{1n}(\tilde{\theta},\tilde{t})$ and $Q_{2n}(\tilde{\theta},\tilde{t})$ at $(\theta_0,0)$ gives

$$\begin{aligned} 0 &= Q_{1n}(\tilde{\theta},\tilde{t}) \\ &= Q_{1n}(\theta_0,0) + \frac{\partial Q_{1n}(\theta_0,0)}{\partial\theta}(\tilde{\theta}-\theta_0) + \frac{\partial Q_{1n}(\theta_0,0)}{\partial t^\tau}(\tilde{t}-0) + o_p(\delta_n) \end{aligned}$$

$$\begin{aligned} 0 &= Q_{2n}(\tilde{\theta},\tilde{t}) \\ &= Q_{2n}(\theta_0,0) + \frac{\partial Q_{2n}(\theta_0,0)}{\partial\theta}(\tilde{\theta}-\theta_0) + \frac{\partial Q_{2n}(\theta_0,0)}{\partial t^\tau}(\tilde{t}-0) + o_p(\delta_n), \end{aligned}$$

where $\delta_n = ||\tilde{\theta}-\theta_0|| + ||\tilde{t}||$. Observing that $Q_{2n}(\theta_0,0) = 0$, we have

$$S_n \begin{pmatrix} \tilde{t} \\ \tilde{\theta}-\theta_0 \end{pmatrix} = \begin{pmatrix} -Q_{1n}(\theta_0,0) + o_p(\delta_n) \\ o_p(\delta_n) \end{pmatrix}, \tag{A1}$$

where

$$S_n = \left.\begin{pmatrix} \frac{\partial Q_{1n}}{\partial t^\tau} & \frac{\partial Q_{1n}}{\partial\theta} \\ \frac{\partial Q_{2n}}{\partial t^\tau} & 0 \end{pmatrix}\right|_{(\theta_0,0)}.$$

Now we solve (A1) for an expression of $\tilde{t}$. By the law of large numbers, as $n \to \infty$

$$\frac{1}{n}\sum_{i=1}^{n}\frac{\partial g_i(\theta)}{\partial\theta} \longrightarrow E\Big(\frac{\partial g(\theta)}{\partial\theta}\Big).$$

Therefore,

$$\frac{\partial g_{n+1}(\theta)}{\partial\theta} = -\frac{a_n}{n}\sum_{i=1}^{n}\frac{\partial g_i(\theta)}{\partial\theta} = o_p(n).$$

Hence applying the law of large numbers again

$$\begin{aligned} \frac{\partial Q_{1n}(\theta,0)}{\partial\theta} &= \frac{1}{n+1}\sum_{i=1}^{n+1}\frac{\partial g_i(\theta)}{\partial\theta} \\ &= \frac{1}{n+1}\sum_{i=1}^{n}\frac{\partial g_i(\theta)}{\partial\theta} + \frac{1}{n+1}\frac{\partial g_{n+1}(\theta)}{\partial\theta} \\ &= E\Big(\frac{\partial g(\theta)}{\partial\theta}\Big) + o_p(1). \end{aligned}$$

Similarly, we can obtain

$$\frac{\partial Q_{2n}}{\partial t^\tau} = E\frac{\partial g(\theta)}{\partial\theta}^\tau + o_p(1) \qquad \text{and} \qquad -\frac{\partial Q_{1n}}{\partial t^\tau} = Eg(\theta)g(\theta)^\tau + o_p(1).$$

Thus as $n \to \infty$

$$S_n \longrightarrow \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & 0 \end{pmatrix} = \left. \begin{pmatrix} -Egg^\tau & E\frac{\partial g}{\partial \theta} \\ E\frac{\partial g}{\partial \theta}^\tau & 0 \end{pmatrix} \right|_{\theta=\theta_0}.$$

We can see that

$$S_n^{-1} \longrightarrow \begin{pmatrix} S_{11}^{-1} + S_{11}^{-1}S_{12}S_{22.1}^{-1}S_{21}S_{11}^{-1} & -S_{11}^{-1}S_{12}S_{22.1}^{-1} \\ -S_{22.1}^{-1}S_{21}S_{11}^{-1} & S_{22.1}^{-1} \end{pmatrix},$$

where $S_{22.1}^{-1} = \left[\left(E\frac{\partial g}{\partial \theta}\right)^\tau (Egg^\tau)^{-1}\left(E\frac{\partial g}{\partial \theta}\right)\right]^{-1}$. Consequently, (A1) can be solved as

$$\begin{pmatrix} \tilde{t} \\ \tilde{\theta} - \theta_0 \end{pmatrix} = S_n^{-1} \begin{pmatrix} -Q_{1n}(\theta_0, 0) + o_p(\delta_n) \\ o_p(\delta_n) \end{pmatrix},$$

which means

$$\begin{aligned} \tilde{\theta} - \theta_0 &= S_{22.1}^{-1}S_{21}S_{11}^{-1}Q_{1n}(\theta_0, 0) + o_p(\delta_n) \\ \tilde{t} &= -(S_{11}^{-1} + S_{11}^{-1}S_{12}S_{22.1}^{-1}S_{21}S_{11}^{-1})Q_{1n}(\theta_0, 0) + o_p(\delta_n). \end{aligned} \tag{A2}$$

Note that by Central Limit Theorem

$$\begin{aligned} Q_{1n}(\theta_0, 0) &= \frac{1}{n+1}\sum_{i=1}^{n+1} g_i(\theta_0) \\ &= \frac{n^{\frac{1}{2}}}{n+1} \cdot n^{-\frac{1}{2}} \sum_{i=1}^{n} g_i(\theta_0) - n^{-\frac{1}{2}} \cdot \frac{a_n}{n+1} \cdot n^{-\frac{1}{2}} \sum_{i=1}^{n} g_i(\theta_0) \\ &= n^{-\frac{1}{2}} \cdot n^{-\frac{1}{2}} \sum_{i=1}^{n} g_i(\theta_0) + o_p(n^{-\frac{1}{2}}), \end{aligned}$$

which implies

$$\sqrt{n}Q_{1n}(\theta_0, 0) \longrightarrow N(0, Egg^\tau) \quad \text{and} \quad Q_{1n} = O_p(n^{-\frac{1}{2}}). \tag{A3}$$

From (A2), we know that

$$\delta_n = ||\tilde{\theta} - \theta_0|| + ||\tilde{t}|| = O_p(n^{-\frac{1}{2}}).$$

Therefore, we have obtained the desired result

$$\tilde{t} = -(S_{11}^{-1} + S_{11}^{-1}S_{12}S_{22.1}^{-1}S_{21}S_{11}^{-1})Q_{1n}(\theta_0, 0) + o_p(n^{-\frac{1}{2}}) \tag{A4}$$

and

$$\tilde{\theta} - \theta_0 = S_{22.1}^{-1}S_{21}S_{11}^{-1}Q_{1n}(\theta_0, 0) + o_p(n^{-\frac{1}{2}}).$$

In particular, we can see that

$$\tilde{t} = O_p(n^{-\frac{1}{2}}) \quad \text{and} \quad \tilde{\theta} - \theta_0 = O_p(n^{-\frac{1}{2}}).$$

Now we are ready to compute $l(\tilde{\theta}) = l(\tilde{\theta}_1, \tilde{\theta}_2)$. Taylor expansion yields

$$
\begin{aligned}
l(\tilde{\theta}_1, \tilde{\theta}_2) &= \sum_{i=1}^{n+1} \log[1 + \tilde{t}^{\tau} g_i(\tilde{\theta})] \\
&= \sum_{i=1}^{n+1} \left(\tilde{t}^{\tau} g_i(\tilde{\theta}) - \frac{1}{2} (\tilde{t}^{\tau} g_i(\tilde{\theta}))^2 \right) + o_p(1) \\
&= \tilde{t}^{\tau} \sum_{i=1}^{n+1} g_i(\tilde{\theta}) - \frac{1}{2} \tilde{t}^{\tau} \left(\sum_{i=1}^{n+1} g_i(\tilde{\theta}) g_i(\tilde{\theta})^{\tau} \right) \tilde{t} + o_p(1).
\end{aligned} \tag{A5}
$$

Note that expanding $g_i(\tilde{\theta})$ at θ_0, we get

$$
g_i(\tilde{\theta}) = g_i(\theta_0) + \frac{\partial g_i(\theta_0)}{\partial \theta} (\tilde{\theta} - \theta_0) + O_p(n^{-1}),
$$

for $i = 1, 2, \ldots, n$.

Hence

$$
\begin{aligned}
\sum_{i=1}^{n} g_i(\tilde{\theta}) &= \sum_{i=1}^{n} g_i(\theta_0) + \sum_{i=1}^{n} \frac{\partial g_i(\theta_0)}{\partial \theta} \cdot (\tilde{\theta} - \theta_0) + O_p(1) \\
&= n Q_{1n}(\theta_0, 0) + n S_{12} S_{22.1}^{-1} S_{21} S_{11}^{-1} Q_{1n}(\theta_0, 0) + o_p(n^{\frac{1}{2}})
\end{aligned}
$$

and

$$
g_{n+1}(\tilde{\theta}) = -\frac{a_n}{n} \sum_{i=1}^{n} g_i(\tilde{\theta}) = o_p(n^{\frac{1}{2}}).
$$

Consequently, we can obtain the first term of (A5) as

$$
\tilde{t}^{\tau} \sum_{i=1}^{n+1} g_i(\tilde{\theta}) = -n\, Q_{1n}(\theta_0, 0)^{\tau} (S_{11}^{-1} + S_{11}^{-1} S_{12} S_{22.1}^{-1} S_{21} S_{11}^{-1}) Q_{1n}(\theta_0, 0) + o_p(1).
$$

Now we calculate the second term of (A5). For $i = 1, 2, \ldots, n$,

$$
g_i(\tilde{\theta}) g_i(\tilde{\theta})^{\tau} = g_i(\theta_0) g_i(\theta_0)^{\tau} + O_p(n^{-\frac{1}{2}}).
$$

Thus

$$
\Sigma_{i=1}^{n} g_i(\tilde{\theta}) g_i(\tilde{\theta})^{\tau} = \sum_{i=1}^{n} g_i(\theta_0) g_i(\theta_0)^{\tau} + O_p(n^{\frac{1}{2}}) = -n S_{11} + O_p(n^{\frac{1}{2}}).
$$

Note that

$$
g_{n+1}(\tilde{\theta}) g_{n+1}(\tilde{\theta})^{\tau} = o_p(n^{\frac{1}{2}}) o_p(n^{\frac{1}{2}}) = o_p(n).
$$

We have

$$
\tilde{t}^{\tau} \left(\sum_{i=1}^{n+1} g_i(\tilde{\theta}) g_i(\tilde{\theta})^{\tau} \right) \tilde{t} = -n\, Q_{1n}(\theta_0, 0)^{\tau} (S_{11}^{-1} + S_{11}^{-1} S_{12} S_{22.1}^{-1} S_{21} S_{11}^{-1}) Q_{1n}(\theta_0, 0) + o_p(1).
$$

Finally, we have

$$
l(\tilde{\theta}_1, \tilde{\theta}_2) = -\frac{n}{2} Q_{1n}(\theta_0, 0)^{\tau} (S_{11}^{-1} + S_{11}^{-1} S_{12} S_{22.1}^{-1} S_{21} S_{11}^{-1}) Q_{1n}(\theta_0, 0) + o_p(1). \tag{A6}
$$

Similarly, we can apply the above process to $l(\theta_1^0, \tilde{\theta}_2^0)$. The procedures are sketched as follows. Let $\tilde{\theta}_2^0$ and $\tilde{t_0} = t(\theta_1^0, \tilde{\theta}_2^0)$ satisfy

$$
Q_{1n}(\theta_1^0, \tilde{\theta}_2^0, \tilde{t_0}) = 0 \quad \text{and} \quad Q_{2n}(\theta_1^0, \tilde{\theta}_2^0, \tilde{t_0}) = 0.
$$

Expanding Q_{1n} and Q_{2n} at $(\theta_1^0, \theta_2^0, 0)$ will produce the linear equations

$$H_n \begin{pmatrix} \tilde{t_0} \\ \tilde{\theta}_2^0 - \theta_2^0 \end{pmatrix} = \begin{pmatrix} -Q_{1n}(\theta_0, 0) + o_p(\delta_n') \\ o_p(\delta_n') \end{pmatrix}, \tag{A7}$$

where $\theta_0 = (\theta_1^0, \theta_2^0)$ is the true value of θ, $\delta_n' = ||\tilde{\theta}_2^0 - \theta_2^0|| + ||\tilde{t_0}||$ and as $n \to \infty$

$$H_n \longrightarrow \begin{pmatrix} H_{11} & H_{12} \\ H_{21} & 0 \end{pmatrix} = \begin{pmatrix} -Egg^\tau & E\dfrac{\partial g}{\partial \theta_2} \\ E\dfrac{\partial g}{\partial \theta_2}^\tau & 0 \end{pmatrix}\Bigg|_{\theta=\theta_0}.$$

Note that $H_{11} = S_{11}$.

Solving (A7) gives us

$$\tilde{t_0} = -(H_{11}^{-1} + H_{11}^{-1} H_{12} H_{22.1}^{-1} H_{21} H_{11}^{-1}) Q_{1n}(\theta_0, 0) + o_p(n^{-\frac{1}{2}}) \tag{A8}$$

and

$$\tilde{\theta}_2^0 - \theta_2^0 = H_{22.1}^{-1} H_{21} H_{11}^{-1} Q_{1n}(\theta_0, 0) + o_p(n^{-\frac{1}{2}}).$$

By Taylor expansion, the above estimations yield

$$l(\theta_1^0, \tilde{\theta}_2^0) = -\frac{1}{2} n Q_{1n}(\theta_0, 0)^\tau (H_{11}^{-1} + H_{11}^{-1} H_{12} H_{22.1}^{-1} H_{21} H_{11}^{-1}) Q_{1n}(\theta_0, 0) + o_p(1). \tag{A9}$$

Using (A9) and (A6), we can write

$$\begin{aligned} W(\theta_1^0) =& 2l(\theta_1^0, \tilde{\theta}_2^0) - 2l(\tilde{\theta}_1, \tilde{\theta}_2) \\ =& [(Egg^\tau)^{-\frac{1}{2}} \sqrt{n} Q_{1n}(\theta_0, 0)]^\tau (A - B) [(Egg^\tau)^{-\frac{1}{2}} \sqrt{n} Q_{1n}(\theta_0, 0)] + o_p(1), \end{aligned}$$

where

$$A = (Egg^\tau)^{-\frac{1}{2}} \left(E\frac{\partial g}{\partial \theta}\right) \left[\left(E\frac{\partial g}{\partial \theta}\right)^\tau (Egg^\tau)^{-1} \left(E\frac{\partial g}{\partial \theta}\right)\right]^{-1} \left(E\frac{\partial g}{\partial \theta}\right)^\tau (Egg^\tau)^{-\frac{1}{2}}$$

$$B = (Egg^\tau)^{-\frac{1}{2}} \left(E\frac{\partial g}{\partial \theta_2}\right) \left[\left(E\frac{\partial g}{\partial \theta_2}\right)^\tau (Egg^\tau)^{-1} \left(E\frac{\partial g}{\partial \theta_2}\right)\right]^{-1} \left(E\frac{\partial g}{\partial \theta_2}\right)^\tau (Egg^\tau)^{-\frac{1}{2}}$$

and all the evaluations related to g are performed at the true value θ_0. By assumption, $E\dfrac{\partial g}{\partial \theta}$ has rank k and Egg^τ is positive definite. Therefore, both A and B are non-negative definite and idempotent. By Lemma 1

$$\begin{aligned} & \left(E\frac{\partial g}{\partial \theta}\right) \left[\left(E\frac{\partial g}{\partial \theta}\right)^\tau (Egg^\tau)^{-1} \left(E\frac{\partial g}{\partial \theta}\right)\right]^{-1} \left(E\frac{\partial g}{\partial \theta}\right)^\tau \\ \geq & \left(E\frac{\partial g}{\partial \theta_1}, E\frac{\partial g}{\partial \theta_2}\right) \begin{pmatrix} 0 & 0 \\ 0 & \left[\left(E\frac{\partial g}{\partial \theta_2}\right)^\tau (Egg^\tau)^{-1} \left(E\frac{\partial g}{\partial \theta_2}\right)\right]^{-1} \end{pmatrix} \begin{pmatrix} E\frac{\partial g}{\partial \theta_1}^\tau \\ E\frac{\partial g}{\partial \theta_2}^\tau \end{pmatrix} \\ = & \left(E\frac{\partial g}{\partial \theta_2}\right) \left[\left(E\frac{\partial g}{\partial \theta_2}\right)^\tau (Egg^\tau)^{-1} \left(E\frac{\partial g}{\partial \theta_2}\right)\right]^{-1} \left(E\frac{\partial g}{\partial \theta_2}\right)^\tau, \end{aligned}$$

which means that $A - B$ is non-negative definite. Thus by Result 2, $A - B$ is also idempotent. From (A3), we can see that $(Egg^\tau)^{-\frac{1}{2}} \sqrt{n} Q_{1n}(\theta_0, 0)$ follows the multivariate standard normal distribution asymptotically. Note that $tr(A) = k$ and $tr(B) = k - q$. We have $tr(A - B) = k - (k - q) = q$. The requirement of Lemma 1 is satisfied, which implies

$$W(\theta_1^0) \xrightarrow{d} \chi_q^2.$$

□

References

1. Sharpe, W.F. Mutual fund performance. *J. Bus.* **1966**, *39*, 119–138. [CrossRef]
2. Jobson, J.D.; Korkie, B.M. Performance hypothesis testing with the sharpe and treynor measures. *J. Financ.* **1981**, *36*, 889–908. [CrossRef]
3. Fama, E. The behavior of stock-market prices. *J. Bus.* **1965**, *38*, 34–105. [CrossRef]
4. Leland, H.E. Beyond mean-variance: Risk and performance measurement in a nonsymmetric world. *Financ. Anal. J.* **1999**, *1*, 27–36. [CrossRef]
5. Agarwal, V.; Naik, N.Y. Risk and portfolio decisions involving hedge funds. *Rev. Financ. Stud.* **2004**, *17*, 63–98. [CrossRef]
6. Ingersoll, J.; Spiegel, M.; Goetzmann, W. Portfolio performance manipulation and manipulation-proof performance measures. *Rev. Financ. Stud.* **2007**, *20*, 1503–1546.
7. Samuelson, P. The fundamental approximation theorem of portfolio analysis in terms of means, variances, and higher moments. *Rev. Econ. Stud.* **1970**, *37*, 537–542. [CrossRef]
8. Scott, R.; Horvath, P. On the direction of preference for moments of higher order than variance. *J. Financ.* **1980**, *35*, 915–919. [CrossRef]
9. Zakamouline, V.; Koekebakker, S. Portfolio performance evaluation with generalized Sharpe ratios: Beyond the mean and variance. *J. Bank. Financ.* **2009**, *33*, 1242–1254. [CrossRef]
10. Pierro, M.D.; Mosevich, J. Effects of skewness and kurtosis on portfolio rankings. *Quant. Financ.* **2011**, *11*, 1449–1453. [CrossRef]
11. Thomas, D.R.; Grunkemeier, G.L. Confidence interval estimation of survival probabilities for censored data. *J. Am. Stat. Assoc.* **1975**, *70*, 865–871. [CrossRef]
12. Owen, A. Empirical likelihood ratio confidence intervals for a single functional. *Biometrika* **1988**, *75*, 237–249. [CrossRef]
13. Owen, A. Empirical likelihood ratio confidence regions. *Ann. Stat.* **1990**, *18*, 90–120. [CrossRef]
14. Hall, P.; La Scala, B. Methodology and algorithms of empirical likelihood. *Inter. Stat. Rev.* **1990**, *58*, 109–127. [CrossRef]
15. Qin, J.; Lawless, J. Empirical likelihood and general estimating equations. *Ann. Stat.* **1994**, *22*, 300–325. [CrossRef]
16. Chen, J.; Variyath, A.M.; Abraham, B. Adjusted empirical likelihood and its properties. *J. Comput. Graph. Stat.* **2008**, *17*, 426–443. [CrossRef]
17. Rao, C.R. *Linear Statistical Inference and Its Applications*; Wiley: New York, NY, USA, 1973.
18. Wilks, S.S. The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses. *Ann. Math. Stat.* **1938**, *9*, 60–62. [CrossRef]
19. Wang, H.J.; Zhu, Z. Empirical likelihood for quantile regression models with longitudinal data. *J. Stat. Plan. Inference* **2011**, *141*, 1603–1615. [CrossRef]
20. Chen, J.; Huang, Y. Finite-sample properties of the adjusted empirical likelihood. *J. Nonparametric Stat.* **2013**, *25*, 147–159. [CrossRef]
21. Mertens, E. Comments on the Correct Variance of Estimated Sharpe Ratios in Lo (2002, FAJ) When Returns Are IID. Research Note. Available online: http://www.elmarmertens.com/research/discussion/soprano01.pdf (accessed on 30 January 2018).
22. Gallier, J. *Geometric Methods and Applications for Computer Science and Engineering*; Springer: New York, NY, USA, 2011.

© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).