UMVUE analysis of two performance evaluation ratios

Pietro Maria Sparago[a]

[a]Department of Statistics, London School of Economics, London, UK

ARTICLE HISTORY

Compiled June 20, 2026

ABSTRACT

We study the uniformly minimum-variance unbiased estimators (UMVUE) of two investment performance ratios in two different modelling settings: (1) the Sharpe ratio (SR) associated to the parameters of a geometric Brownian motion model, and (2) the information ratio (IR) associated to a factor model. In (1), in particular, we highlight how the estimation of the Sharpe ratio with log-returns from discrete sampling must take into account the variance (or Itô) correction in the log-price dynamics. In (2), we also derive the full distribution of the IR UMVUE, which is of interest in its own right. All methods in this note follow well-established foundational arguments of mathematical statistics and unbiased estimation; we will cite [1] often, as a graduate-level reference.

1. UMVUE of the continuous-time Sharpe ratio in GBM dynamics

Suppose that the price process $S = (S_t)_{t\geq 0}$ follows a geometric Brownian motion with strictly positive volatility $\sigma > 0$ and initial condition $S_0 = s_0 > 0$ and we want to estimate the Sharpe ratio (SR) of the continuous-time dynamics $dS_t/S_t = \mu dt + \sigma dW_t$ (i.e. the Sharpe ratio of the infinitesimal return):

$$\mathrm{SR} := \frac{\mu}{\sigma}$$

The process S is sampled at equally-spaced observation times $\{0, \Delta t, 2\Delta t, ..., k\Delta t, ...\}$ where $\Delta t > 0$. A standard argument (indeed: apply Itô's formula to $f(S_t) = \ln S_t$) yields log-returns $X_k := \ln(S_{k\Delta t}/S_{(k-1)\Delta t}),\ k \in \mathbb{N}$ with $X_k \sim \mathcal{N}((\mu - \sigma^2/2)\Delta t, \sigma^2\Delta t),\ \mathrm{IID}$. The variance correction $-\sigma^2\Delta t/2$ makes the standard Sharpe ratio UMVUE (see e.g. ([1]; Ex. 3.4 p. 164)) imprecise, yielding only the unbiased estimator of $(\mu - \sigma^2/2)/\sigma$. For a random sample of $n \geq 4$ log-returns, define the sample mean $\overline{X}_n$ and the sample variance S_n^2:

$$\overline{X}_n := \frac{1}{n}\sum_{1\leq k\leq n} X_k \implies \overline{X}_n \sim \mathcal{N}\left((\mu - \sigma^2/2)\Delta t, \frac{\sigma^2\Delta t}{n}\right)$$

$$S_n^2 := \frac{1}{n-1}\sum_{k\leq n}(X_k - \overline{X}_n)^2 \implies \frac{(n-1)S_n^2}{\sigma^2\Delta t} \sim \chi^2_{n-1}$$

CONTACT P. M. Email: p.sparago@lse.ac.uk

Then, $\overline{X}_n$ and S_n^2 are independent. Indeed, define $Y_k := X_k + \sigma^2\Delta t/2$. The sample mean $\overline{Y}_n$ is complete and sufficient for μ when σ^2 is known (because $\overline{Y}_n \sim \mathcal{N}(\mu\Delta t, \sigma^2\Delta t/n)$, see e.g. ([1]; Ex. 2.18 p. 112)), while S_n^2 is ancillary for σ^2 fixed: then, $\overline{Y}_n$ is independent of S_n^2 by Basu's theorem, hence also $\overline{X}_n = \overline{Y}_n - \sigma^2\Delta t/2$. We can now compute the expectation of the natural candidate:

$$E\left[\frac{\overline{X}_n + S_n^2/2}{S_n}\right] = \Delta t\left(\mu - \frac{\sigma^2}{2}\right)E[S_n^{-1}] + \frac{1}{2}E[S_n]$$

We have:

$$\begin{aligned}
E[S_n^{-1}] &= \frac{\sqrt{n-1}}{\sigma\sqrt{\Delta t}}E\left[\left(\frac{(n-1)S_n^2}{\sigma^2\Delta t}\right)^{-1/2}\right] \\
&= \frac{\sqrt{n-1}}{\sigma\sqrt{\Delta t}}\int_0^\infty x^{-1/2}\frac{x^{(n-1)/2-1}e^{-x/2}}{2^{(n-1)/2}\Gamma((n-1)/2)}dx \\
&= \frac{\sqrt{n-1}}{\sigma\sqrt{\Delta t}}\int_0^\infty \frac{x^{(n-2)/2-1}e^{-x/2}}{2^{(n-1)/2}\Gamma((n-1)/2)}dx \\
&= \frac{\sqrt{n-1}}{\sigma\sqrt{\Delta t}}\frac{2^{-1/2}\Gamma((n-2)/2)}{\Gamma((n-1)/2)} \\
&= \frac{1}{\sigma\sqrt{\Delta t}}\frac{2^{-1/2}\Gamma((n-2)/2)}{(n-1)^{-1/2}\Gamma((n-1)/2)} \\
E[S_n] &= \sigma\sqrt{\Delta t}\frac{(n-1)^{-1/2}\Gamma(n/2)}{2^{-1/2}\Gamma((n-1)/2)}
\end{aligned}$$

So that

$$E\left[\frac{\overline{X}_n/C_n + S_n^2/(2D_n)}{S_n}\right] = \sqrt{\Delta t}\left(\mu - \frac{\sigma^2}{2}\right)\frac{1}{\sigma} + \frac{\sigma\sqrt{\Delta t}}{2} = \frac{\mu}{\sigma}\sqrt{\Delta t}$$

where

$$C_n := \frac{2^{-1/2}\Gamma((n-2)/2)}{(n-1)^{-1/2}\Gamma((n-1)/2)}, \quad D_n := \frac{(n-1)^{-1/2}\Gamma(n/2)}{2^{-1/2}\Gamma((n-1)/2)}$$

So we can conclude that the UMVUE of the Sharpe ratio μ/σ is

$$\boxed{\hat{\mathrm{SR}}(X_1, ..., X_n) := \frac{1}{S_n\sqrt{\Delta t}}\left(\frac{\overline{X}_n}{C_n} + \frac{S_n^2}{2D_n}\right), \quad n \geq 4}$$

since $(\overline{X}_n, S_n^2)$ is complete and sufficient, by Lehmann-Scheffé ([1]; Th. 3.1 p. 162). Note that $n \geq 4$ is needed for $V[\hat{\mathrm{SR}}]$ to be finite.

2. UMVUE of the information ratio in a linear factor model

Consider the factor model specification $Y_k = \alpha + \beta X_k + \varepsilon_k$, $\varepsilon_k \sim \mathcal{N}(0, \sigma^2)$, IID where $X_1, X_2, ..., X_n$ for $n \geq 5$ are known covariates and $\sigma > 0$. The information ratio (IR)

of the model is defined as

$$\mathrm{IR} := \frac{\alpha}{\sigma}$$

Define the covariates sample variance $S^2_{X,n}$ (assumed strictly positive), the regression estimators $\hat{\alpha}_n, \hat{\beta}_n$ and the (scaled) residual sum-of-squares $\tilde{S}^2_n$:

$$\begin{aligned}
S^2_{X,n} &:= \frac{1}{n-1} \sum_{1\le k\le n} (X_k - \overline{X}_n)^2 \\
\hat{\beta}_n &:= \frac{1}{(n-1)S^2_{X,n}} \sum_{1\le k\le n} (X_k - \overline{X}_n)(Y_k - \overline{Y}_n) \\
\hat{\alpha}_n &:= \overline{Y}_n - \hat{\beta}_n \overline{X}_n \\
\tilde{S}^2_n &:= \frac{1}{n-2} \sum_{1\le k\le n} (Y_k - \hat{\beta}_n X_k - \hat{\alpha}_n)^2 \implies \frac{(n-2)\tilde{S}^2_n}{\sigma^2} \sim \chi^2_{n-2}
\end{aligned}$$

The regression estimators are unbiased estimators of α, β and it is well-known that $\hat{\alpha}_n$ (and $\hat{\beta}_n$ as well) is independent of $\tilde{S}^2_n$ (see e.g. ([1]; Th. 3.8 p. 188)). We get $\sqrt{n-2}\tilde{S}_n/\sigma \sim \chi_{n-2}$, hence we can evaluate the expectation of the natural candidate:

$$\begin{aligned}
E\left[\frac{\hat{\alpha}_n}{\tilde{S}_n}\right] &= \alpha E[\tilde{S}^{-1}_n] \\
&= \frac{\alpha\sqrt{n-2}}{\sigma} \int_0^\infty x^{-1} \left(\frac{x^{(n-2)-1} e^{-x^2/2}}{2^{(n-2)/2-1}\Gamma((n-2)/2)} \right) dx \\
&= \frac{\alpha\sqrt{n-2}}{\sigma} \int_0^\infty \left(\frac{x^{(n-3)-1} e^{-x^2/2}}{2^{(n-2)/2-1}\Gamma((n-2)/2)} \right) dx \\
&= \frac{2^{(n-3)/2-1}\Gamma((n-3)/2)}{2^{(n-2)/2-1}\Gamma((n-2)/2)} \frac{\alpha\sqrt{n-2}}{\sigma} \\
&= \sqrt{\frac{n-2}{2}} \frac{\Gamma((n-3)/2)}{\Gamma((n-2)/2)} \frac{\alpha}{\sigma}
\end{aligned}$$

So we conclude that the UMVUE of the information ratio α/σ is

$$\boxed{\hat{\mathrm{IR}}(X_1, ..., X_n; Y_1, ..., Y_n) = \sqrt{\frac{2}{n-2}} \frac{\Gamma((n-2)/2)}{\Gamma((n-3)/2)} \frac{\hat{\alpha}_n}{\tilde{S}_n}, \quad n \ge 5}$$

again by Lehmann-Scheffé ([1]; Th. 3.1 p. 162). Note that an entirely equivalent argument holds for the UMVUE of β/σ. The restriction $n \ge 5$ is needed for $V[\hat{\mathrm{IR}}]$ to be finite. We shall now derive the finite-sample distribution of $\hat{\mathrm{IR}}$. Define

$$\begin{aligned}
F_n &:= \sqrt{\frac{n-2}{2}} \frac{\Gamma((n-3)/2)}{\Gamma((n-2)/2)} \\
\gamma^2_n &:= \frac{1}{F^2_n} \left(\frac{1}{n} + \frac{\overline{X}^2_n}{(n-1)S^2_{X,n}} \right)
\end{aligned}$$

We can then rewrite the UMVUE as

$$\hat{\mathrm{IR}}(X_1, ..., X_n; Y_1, ..., Y_n) = \frac{\hat{\alpha}_n}{\sigma F_n} \Bigg/ \left(\frac{1}{\sqrt{n-2}} \frac{\tilde{S}_n \sqrt{n-2}}{\sigma} \right)$$

where

$$\frac{\hat{\alpha}_n}{\sigma F_n} \sim \mathcal{N}\left(\frac{\alpha}{\sigma F_n}, \gamma_n^2 \right)$$

So we get the noncentral t-distribution

$$\frac{\hat{\mathrm{IR}}(X_1, ..., X_n; Y_1, ..., Y_n)}{\gamma_n} = \frac{\hat{\alpha}_n / \tilde{S}_n}{\sqrt{\frac{1}{n} + \frac{\overline{X}_n^2}{(n-1)S_{X,n}^2}}} \sim t\left(n-2, \frac{\alpha}{\sigma \sqrt{\frac{1}{n} + \frac{\overline{X}_n^2}{(n-1)S_{X,n}^2}}} \right)$$

Which ultimately implies

$$\boxed{\hat{\mathrm{IR}}(X_1, ..., X_n; Y_1, ..., Y_n) \sim \frac{\sqrt{\frac{1}{n} + \frac{\overline{X}_n^2}{(n-1)S_{X,n}^2}}}{F_n} \cdot t\left(n-2, \frac{\alpha}{\sigma \sqrt{\frac{1}{n} + \frac{\overline{X}_n^2}{(n-1)S_{X,n}^2}}} \right)}$$

that is, the UMVUE has a scaled noncentral t-distribution. Sanity check: for $n \geq 5$

$$E[\hat{\mathrm{IR}}(X_1, ..., X_n; Y_1, ..., Y_n)] = \frac{\sqrt{\frac{1}{n} + \frac{\overline{X}_n^2}{(n-1)S_{X,n}^2}}}{F_n} F_n \frac{\alpha}{\sigma \sqrt{\frac{1}{n} + \frac{\overline{X}_n^2}{(n-1)S_{X,n}^2}}} = \frac{\alpha}{\sigma}$$

which is consistent with what we have shown above.

References

[1] Shao, J. (2003). *Mathematical Statistics (Second Edition)*. Springer-Verlag New York.