Inference on Achieved Signal Noise Ratio

Steven E. Pav *

May 16, 2020

Abstract

We describe a procedure to perform approximate inference on the achieved signal-noise ratio of the Markowitz portfolio under Gaussian *i.i.d.* returns. The procedure relies on a statistic similar to the Sharpe Ratio Information Criterion. [10] Testing indicates the procedure is somewhat conservative, but otherwise works well for reasonable values of sample and asset universe sizes.

1 Introduction

For a universe of k assets, we consider the portfolio optimization problem

$$\max_{\boldsymbol{\nu}} \frac{\boldsymbol{\nu}^\top \boldsymbol{\mu}}{\sqrt{\boldsymbol{\nu}^\top \Sigma \boldsymbol{\nu}}}. \tag{1}$$

Here $\boldsymbol{\mu}$ is the expected return and Σ is the covariance of returns. This problem is solved by the Markowitz portfolio, defined as

$$\boldsymbol{\nu}_* =_{\mathrm{df}} \Sigma^{-1} \boldsymbol{\mu}, \tag{2}$$

and any positive multiple thereof.

In practice the parameters $\boldsymbol{\mu}$ and Σ are unknown and must be estimated from the data. The estimation of parameters is known to deterioriate the quality of the portfolio. [9] The signal-noise ratio of the Markowitz portfolio, its mean divided by its volatility, is subject to a fundamental bound. [12, 2] While inference on the population parameters follows from classical statistics via the connection to Hotelling's T^2, little is known about performing inference on the signal-noise ratio achieved by the Markowitz portfolio. Paulsen and Söhl described the *Sharpe Ratio Information Criterion* (SRIC), which is an approximately unbiased estimator for this quantity. [10] Some asymptotic confidence intervals have also been described, but these require unreasonably large sample sizes. [11] Here we fill this gap, describing confidence intervals very similar to the SRIC and using the same approximation. Practical construction of these bounds requires one to estimate the population effect size. In practice this causes the confidence intervals to be slightly conservative.

*steven@gilgamath.com The code to build this document is available at www.github.com/shabbychef/snrinf. This revision was built from commit 55f8176fbc33bcf6b03f02174afb0d7e394105cf of that repo.

2 The Procedure

Assume you observe returns on k assets, which are independently drawn from a Gaussian distribution $x_t \sim \mathcal{N}(\boldsymbol{\mu}, \Sigma)$. The population Markowitz portfolio is $\boldsymbol{\nu}_* = \Sigma^{-1}\boldsymbol{\mu}$. The signal-noise ratio of this portfolio is $\zeta_* = \sqrt{\boldsymbol{\mu}^\top \Sigma^{-1}\boldsymbol{\mu}}$. Given n observations of returns, one typically estimates the population parameters via

$$\hat{\boldsymbol{\mu}} = \frac{1}{n}\sum_{1\le t\le n} x_t, \tag{3}$$

$$\hat{\Sigma} = \frac{1}{n-1}\sum_{1\le t\le n} x_t {x_t}^\top - \frac{n}{n-1}\hat{\boldsymbol{\mu}}\hat{\boldsymbol{\mu}}^\top. \tag{4}$$

The (sample) Markowitz portfolio is $\hat{\boldsymbol{\nu}}_* = \hat{\Sigma}^{-1}\hat{\boldsymbol{\mu}}$. The achieved signal-noise ratio of $\hat{\boldsymbol{\nu}}_*$ is defined as

$$\zeta_a =_{\mathrm{df}} \frac{\boldsymbol{\mu}^\top \hat{\boldsymbol{\nu}}_*}{\sqrt{{\hat{\boldsymbol{\nu}}_*}^\top \Sigma \hat{\boldsymbol{\nu}}_*}}. \tag{5}$$

It is an unobservable random quantity that we wish to perform inference on.

The Sharpe ratio of $\hat{\boldsymbol{\nu}}_*$ is defined as

$$\hat{\zeta}_* =_{\mathrm{df}} \frac{\hat{\boldsymbol{\mu}}^\top \hat{\boldsymbol{\nu}}_*}{\sqrt{{\hat{\boldsymbol{\nu}}_*}^\top \hat{\Sigma} \hat{\boldsymbol{\nu}}_*}} = \sqrt{\hat{\boldsymbol{\mu}}^\top \hat{\Sigma}^{-1}\hat{\boldsymbol{\mu}}}. \tag{6}$$

We note that $T^2 = n\hat{\zeta}_*^2$ is the familiar Hotelling's statistic, which is usually prescribed to perform inference on $\boldsymbol{\mu}$, but can be used to perform inference on ζ_*^2. [1, 13]

The Sharpe Ratio Information Criterion is defined as [10]

$$SRIC =_{\mathrm{df}} \hat{\zeta}_* - \frac{k-1}{n\hat{\zeta}_*}. \tag{7}$$

Under the simplifying approximation

$$\hat{\Sigma} \approx \Sigma, \tag{8}$$

the SRIC is unbiased for the achieved signal-noise ratio:

$$\mathrm{E}\left[SRIC\right] = \mathrm{E}\left[\zeta_a\right]. \tag{9}$$

Note this only holds for $k > 1$, but it is simple to express $\mathrm{E}\left[\zeta_a\right]$ when $k = 1$.

Inspired by the SRIC, we seek a constant c_α such that

$$\Pr\left\{\zeta_a \le \hat{\zeta}_* - \frac{c_\alpha}{n\hat{\zeta}_*}\right\} = \alpha. \tag{10}$$

Under Approximation 8,

$$\zeta_a \approx \frac{\boldsymbol{\mu}^\top \Sigma^{-1}\hat{\boldsymbol{\mu}}}{\sqrt{\left(\Sigma^{-1}\hat{\boldsymbol{\mu}}\right)^\top \Sigma \left(\Sigma^{-1}\hat{\boldsymbol{\mu}}\right)}} = \frac{\boldsymbol{\mu}^\top \Sigma^{-1}\hat{\boldsymbol{\mu}}}{\sqrt{\hat{\boldsymbol{\mu}}^\top \Sigma^{-1}\hat{\boldsymbol{\mu}}}} = \frac{\boldsymbol{\mu}^\top \Sigma^{-1}\hat{\boldsymbol{\mu}}}{\hat{\zeta}_*}. \tag{11}$$

Under the approximation we also have $\hat{\zeta}_*^2 \approx \hat{\boldsymbol{\mu}}^\top \Sigma^{-1} \hat{\boldsymbol{\mu}}$. We note that for Gaussian returns, we can write

$$\hat{\boldsymbol{\mu}} = \boldsymbol{\mu} + \frac{1}{\sqrt{n}} \Sigma^{1/2} \boldsymbol{z},$$

where $\boldsymbol{z} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$. Thus

$$\begin{aligned}
\hat{\zeta}_*^2 - \zeta_a \hat{\zeta}_* &\approx \hat{\boldsymbol{\mu}}^\top \Sigma^{-1} \hat{\boldsymbol{\mu}} - \boldsymbol{\mu}^\top \Sigma^{-1} \hat{\boldsymbol{\mu}}, \\
&= (\hat{\boldsymbol{\mu}} - \boldsymbol{\mu})^\top \Sigma^{-1} \hat{\boldsymbol{\mu}}, \\
&= \frac{1}{n} \boldsymbol{z}^\top \Sigma^{\top/2} \Sigma^{-1} \left(\sqrt{n} \boldsymbol{\mu} + \Sigma^{1/2} \boldsymbol{z} \right), \\
&= \frac{1}{n} \boldsymbol{z}^\top \left(\sqrt{n} \Sigma^{-1/2} \boldsymbol{\mu} + \boldsymbol{z} \right), \\
&= \frac{1}{n} \left(\frac{1}{2} \sqrt{n} \Sigma^{-1/2} \boldsymbol{\mu} + \boldsymbol{z} - \frac{1}{2} \sqrt{n} \Sigma^{-1/2} \boldsymbol{\mu} \right)^\top \left(\frac{1}{2} \sqrt{n} \Sigma^{-1/2} \boldsymbol{\mu} + \boldsymbol{z} + \frac{1}{2} \sqrt{n} \Sigma^{-1/2} \boldsymbol{\mu} \right), \\
&= \frac{1}{n} \left(\left\| \frac{1}{2} \sqrt{n} \Sigma^{-1/2} \boldsymbol{\mu} + \boldsymbol{z} \right\|_2^2 - \left\| \frac{1}{2} \sqrt{n} \Sigma^{-1/2} \boldsymbol{\mu} \right\|_2^2 \right), \\
&\sim \frac{1}{n} \left(\chi^2 \left(k, \frac{n \zeta_*^2}{4} \right) - \frac{n \zeta_*^2}{4} \right).
\end{aligned}$$

Now because

$$\zeta_a \leq \hat{\zeta}_* - \frac{c}{n \hat{\zeta}_*} \Leftrightarrow c \leq n \left(\hat{\zeta}_*^2 - \zeta_a \hat{\zeta}_* \right),$$

If we want this condition to hold with probability α we should set

$$c_\alpha = {\chi^2}_{1-\alpha} \left(k, \frac{n \zeta_*^2}{4} \right) - \frac{n \zeta_*^2}{4}, \tag{12}$$

where ${\chi^2}_q(v, \delta)$ is the q quantile of the non-central chi-square distribution with v degrees of freedom and non-centrality parameter δ.

Checking Coverage Before proceeding, we check whether use of Approximation 8 leads to a degradation in coverage of a confidence interval implied by Inequality 10. We draw n days of returns from the k-variate normal distribution. For a fixed value of ζ_*, we perform 1000000 simulations of computing ζ_a and $\hat{\zeta}_*^2$, computing a one-sided confidence bound and measuring the empirical rate of type I errors. We then let n vary from 50 to 102,400 days; we let k vary from 2 to 16; we let ζ_* vary from $0.5\,\mathrm{yr}^{-1/2}$ to $2\,\mathrm{yr}^{-1/2}$, where we assume 252 days per year. We compute the lower confidence limit on ζ_a using knowledge of the actual ζ_* to construct c_α. For practical inference this would have to be estimated, but here we are only testing conditions for which the approximation $\hat{\Sigma} \approx \Sigma$ is close enough for purposes of inference.

In Figure 1 we plot the empirical type I rate at the nominal 0.05 level of the confidence bound. The main takeaway from this experiment is that the bound gives near-nominal coverage when $n \geq 100k$ or so.

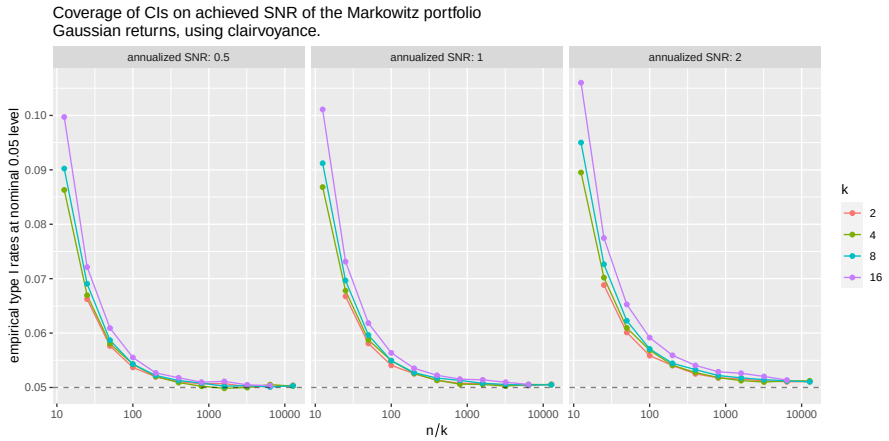

Figure 1: The empirical type I rate, over 1,000,000 simulations, of a one-sided confidence bound for ζ_a are shown for a nominal type I rate of 0.05. The daily returns are drawn from multivariate normal distribution with varying ζ_*, n, and k. Type I rates are plotted versus n/k to indicate the requisite aspect ratio to achieve near nominal coverage.

2.1 Practical Inference

One can construct one- or two-sided confidence intervals from Inequality 10 when ζ_* is known. However, it is unknown in practice, and the constant c_α is sufficiently sensitive to it. To practically perform inference, there are two obvious routes: one is to jointly perform inference on ζ_* on ζ_a; the other is to estimate ζ_* and plug it in when constructing c_α.

For the joint estimation procedure, for some $q \in (0,1)$, construct a $q\alpha$ upper bound on ζ_*. That confidence bound can be described implicitly via the connection to the non-central F distribution: to find the one-sided confidence intervals $[0, \zeta_u]$ with coverage $1 - q\alpha$, find

$$\zeta_u = \min\left\{ z \,\middle|\, z \geq 0,\ \alpha/2 \geq F_f\left(\left(\frac{n(n-k)}{k(n-1)}\right)\hat{\zeta}_*^2; k, n-k, nz^2\right)\right\}, \tag{13}$$

where $F_f\left(x; \nu_1, \nu_2, \delta\right)$ is the CDF of the non-central F-distribution with non-centrality parameter δ and ν_1 and ν_2 degrees of freedom. This method requires computational inversion of the CDF function. Then compute

$$c = \max\left\{ {\chi^2}_{1-(1-q)\alpha}\left(k, \frac{n\zeta^2}{4}\right) - \frac{n\zeta^2}{4} \,\middle|\, 0 \leq \zeta_u \right\}.$$

The bound $\hat{\zeta}_* - \frac{c}{n\hat{\zeta}_*}$ then should have type I rate at most α. However, since this is a joint confidence bound the bound on ζ_a will be somewhat conservative.

Another approach, which does not have guaranteed coverage, is to estimate ζ_* from the data, and plug in that value in the computation of c_α. We can perform this estimation using standard techniques, again via the connection of Hotelling's T^2 to the F distribution. Kubokawa, Robert and Saleh described improved methods for estimating the non-centrality parameter given an observation of a non-central F statistic. [8]. They described the following estimators

for the non-centrality parameter, which is ζ_*^2 in our case:

$$\begin{aligned}
\delta_0 &= \frac{(n-k-2)}{n-1}\hat{\zeta}_*^2 - \frac{k}{n}, \\
\delta_1 &= max\left(\delta_0, 0\right), \\
\delta_2 &= max\left(\delta_0, \frac{2}{k+2}\left(\delta_0 + \frac{k}{n}\right)\right).
\end{aligned} \tag{14}$$

They note that δ_0 is the Uniform Minimum Variance Unbiased Estimator (UMVUE) of ζ_*^2. However, it can be negative. The estimators δ_1, δ_2 are non-negative, and dominate δ_0 in having lower expected squared error. Thus the suggested procedure is to compute

$$c = {\chi^2}_{1-\alpha}\left(k, \frac{n\delta_2}{4}\right) - \frac{n\delta_2}{4},$$

then use the bound $\hat{\zeta}_* - \frac{c}{n\hat{\zeta}_*}$. In practice this bound seems to give slightly less conservative coverage than the joint bound described above. It is not clear how to find a coverage guarantee for this bound. The quantities $\hat{\zeta}_*^2$ and ζ_a are not independent, and their asymptotic correlation is $\mathcal{O}\left(n^{-1/2}\right)$, which is only slowly shrinking. [11]

Feasible CI Coverage We reconsider the experiments above but compute feasible confidence bounds. We use both the simultaneous CI approach with $q = 0.25$; and plug in $\zeta_* = \sqrt{\delta_2}$ to construct the bound. In Figure 2, we plot the empirical type I rate for both of these bounds versus n, with facets for ζ_* and k. We see that the δ_2 plug-in estimator has coverage closer to the nominal 0.05 rate. Both bounds have issues when n/k is not sufficiently large, a problem stemming from the poor quality of the approximation $\hat{\Sigma} \approx \hat{\Sigma}$, and which was seen above. However, here we see closer to nominal coverage for larger k for both methods. It is not clear how the coverage will behave for larger n/k, though that seems like an unlikely problem in practice.

2.2 Hedged Portfolios

Now we generalize the portfolio problem of Equation 1 to add a hedging constraint. So consider the constrained portfolio optimization problem on k assets,

$$\max_{\substack{\mathsf{G}\Sigma\boldsymbol{\nu}=\mathbf{0}, \\ \boldsymbol{\nu}^\top\Sigma\boldsymbol{\nu}\leq R^2}} \frac{\boldsymbol{\nu}^\top\boldsymbol{\mu} - r_0}{\sqrt{\boldsymbol{\nu}^\top\Sigma\boldsymbol{\nu}}}, \tag{15}$$

where G is an $k_g \times k$ matrix of rank k_g, and, as previously, $\boldsymbol{\mu}$, Σ are the mean vector and covariance matrix, r_0 is the risk-free rate, and $R > 0$ is a risk 'budget'. We can interpret the G constraint as stating that the covariance of the returns of a feasible portfolio with the returns of a portfolio whose weights are in a given row of G shall equal zero. In the garden variety application of this problem, G consists of k_g rows of the identity matrix; in this case, feasible portfolios are *hedged* with respect to the k_g assets selected by G (although they may hold some position in the hedged assets). We use "hedged" to mean a portfolio with zero covariance against some other portfolio(s).

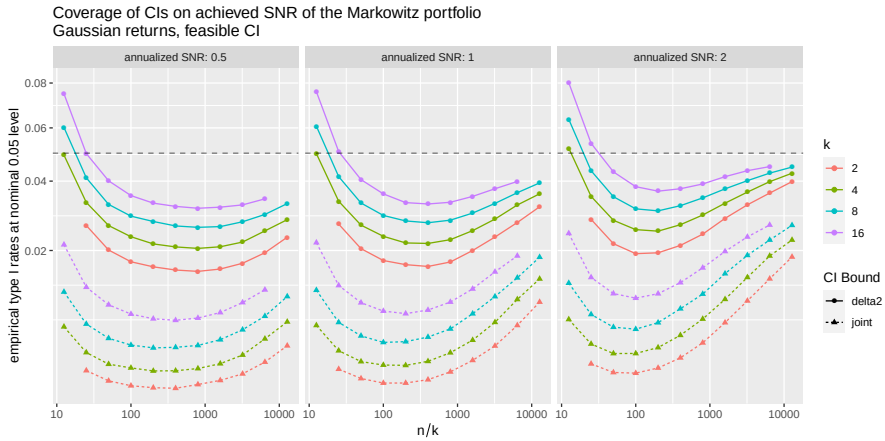

Figure 2: The empirical type I rate, over 1,000,000 simulations, of two feasible one-sided confidence bounds for ζ_a are shown for a nominal type I rate of 0.05. The daily returns are drawn from multivariate normal distribution with varying ζ_*, n, and p. The y axis is drawn in square root scale to show detail.

The solution to this problem, via the Lagrange multiplier technique, is

$$\boldsymbol{\nu}_{*,\mathsf{I}\setminus\mathsf{G}} = c\left(\Sigma^{-1}\boldsymbol{\mu} - \mathsf{G}^\top\left(\mathsf{G}\Sigma\mathsf{G}^\top\right)^{-1}\mathsf{G}\boldsymbol{\mu}\right).$$

When $r_0 > 0$, the unique solution is found by setting c so that the risk budget is an equality. Note that, up to scaling, $\Sigma^{-1}\boldsymbol{\mu}$ is the unconstrained optimal portfolio, and thus the imposition of the G constraint only changes the unconstrained portfolio in assets corresponding to columns of G containing non-zero elements. In the garden variety application where G is a single row of the identity matrix, the imposition of the constraint only changes the holdings in the asset to be hedged (modulo changes in the leading constant to satisfy the risk budget).

The squared signal-noise ratio of the optimal portfolio we write as

$$\Delta_{\mathsf{I}\setminus\mathsf{G}}\zeta_*^2 =_{\mathrm{df}} \boldsymbol{\mu}^\top\Sigma^{-1}\boldsymbol{\mu} - (\mathsf{G}\boldsymbol{\mu})^\top\left(\mathsf{G}\Sigma\mathsf{G}^\top\right)^{-1}(\mathsf{G}\boldsymbol{\mu}). \tag{16}$$

The sample optimal portfolio is given by

$$\hat{\boldsymbol{\nu}}_{*,\mathsf{I}\setminus\mathsf{G}} = c\left(\hat{\Sigma}^{-1}\hat{\boldsymbol{\mu}} - \mathsf{G}^\top\left(\mathsf{G}\hat{\Sigma}\mathsf{G}^\top\right)^{-1}\mathsf{G}\hat{\boldsymbol{\mu}}\right).$$

The squared Sharpe ratio of this portfolio is

$$\Delta_{\mathsf{I}\setminus\mathsf{G}}\hat{\zeta}_*^2 = \hat{\boldsymbol{\mu}}^\top\hat{\Sigma}^{-1}\hat{\boldsymbol{\mu}} - (\mathsf{G}\hat{\boldsymbol{\mu}})^\top\left(\mathsf{G}\hat{\Sigma}\mathsf{G}^\top\right)^{-1}(\mathsf{G}\hat{\boldsymbol{\mu}}). \tag{17}$$

The achieved signal-noise ratio of this portfolio is

$$\zeta_a = \frac{\boldsymbol{\mu}^\top\hat{\boldsymbol{\nu}}_{*,\mathsf{I}\setminus\mathsf{G}}}{\sqrt{\hat{\boldsymbol{\nu}}_{*,\mathsf{I}\setminus\mathsf{G}}{}^\top\Sigma\hat{\boldsymbol{\nu}}_{*,\mathsf{I}\setminus\mathsf{G}}}}. \tag{18}$$

Define:

$$\hat{\zeta}_{*,\mathsf{G}}^2 =_{\mathrm{df}} (\mathsf{G}\hat{\boldsymbol{\mu}})^\top\left(\mathsf{G}\hat{\Sigma}\mathsf{G}^\top\right)^{-1}(\mathsf{G}\hat{\boldsymbol{\mu}}).$$

Giri showed that conditional on observing $\hat{\zeta}^2_{*,\mathsf{G}}$,

$$\frac{n}{n-1}\frac{n-k}{k-k_g}\frac{\Delta_{\mathsf{I}\setminus\mathsf{G}}\hat{\zeta}^2_*}{1+\frac{n}{n-1}\hat{\zeta}^2_{*,\mathsf{G}}}\Big|\,\hat{\zeta}^2_{*,\mathsf{G}} \sim F\left(k-k_g, n-k, \frac{n}{1+\frac{n}{n-1}\hat{\zeta}^2_{*,\mathsf{G}}}\Delta_{\mathsf{I}\setminus\mathsf{G}}\zeta^2_*\right), \tag{19}$$

where $F(v_1, v_2, \delta)$ is the non-central F-distribution with v_1, v_2 degrees of freedom and non-centrality parameter δ. [6, 13]

Now we apply Approximation 8, and complete the square as we did in the unhedged case, to find that

$$\begin{aligned}\Delta_{\mathsf{I}\setminus\mathsf{G}}\zeta^2_* - \zeta_a\sqrt{\Delta_{\mathsf{I}\setminus\mathsf{G}}\zeta^2_*} &\approx (\hat{\boldsymbol{\mu}}-\boldsymbol{\mu})^\top\Sigma^{-1}\hat{\boldsymbol{\mu}} - (\hat{\boldsymbol{\mu}}-\boldsymbol{\mu})^\top\mathsf{G}^\top\left(\mathsf{G}\Sigma\mathsf{G}^\top\right)^{-1}\mathsf{G}\hat{\boldsymbol{\mu}},\\ &= \frac{1}{n}\left[\left(\frac{1}{2}\sqrt{n}\Sigma^{-1/2}\boldsymbol{\mu}+\boldsymbol{z}\right)^\top\left(\mathsf{I}-\Sigma^{\top/2}\mathsf{G}^\top\left(\mathsf{G}\Sigma\mathsf{G}^\top\right)^{-1}\mathsf{G}\Sigma^{1/2}\right)\left(\frac{1}{2}\sqrt{n}\Sigma^{-1/2}\boldsymbol{\mu}+\boldsymbol{z}\right)\right]\\ &\quad-\frac{1}{4n}\left[\left(\sqrt{n}\Sigma^{-1/2}\boldsymbol{\mu}\right)^\top\left(\mathsf{I}-\Sigma^{\top/2}\mathsf{G}^\top\left(\mathsf{G}\Sigma\mathsf{G}^\top\right)^{-1}\mathsf{G}\Sigma^{1/2}\right)\left(\sqrt{n}\Sigma^{-1/2}\boldsymbol{\mu}\right)\right],\end{aligned}$$

where $\boldsymbol{z}\sim\mathcal{N}(\mathbf{0},\mathsf{I})$. Now note that the matrix

$$\mathsf{A} = \mathsf{I} - \Sigma^{\top/2}\mathsf{G}^\top\left(\mathsf{G}\Sigma\mathsf{G}^\top\right)^{-1}\mathsf{G}\Sigma^{1/2}$$

is idempotent with rank $k-k_g$. Thus a quadratic form in A follows a non-central χ^2 distribution[1] with degrees of freedom equal to the rank of A. [7, Theorem 2] Thus

$$\Delta_{\mathsf{I}\setminus\mathsf{G}}\zeta^2_* - \zeta_a\sqrt{\Delta_{\mathsf{I}\setminus\mathsf{G}}\zeta^2_*} \sim \frac{1}{n}\left(\chi^2\left(k-k_g, \frac{n\Delta_{\mathsf{I}\setminus\mathsf{G}}\zeta^2_*}{4}\right) - \frac{n\Delta_{\mathsf{I}\setminus\mathsf{G}}\zeta^2_*}{4}\right).$$

Then, as in the unhedged case, we have

$$\Pr\left\{\zeta_a \le \sqrt{\Delta_{\mathsf{I}\setminus\mathsf{G}}\hat{\zeta}^2_*} - \frac{c_\alpha}{n\sqrt{\Delta_{\mathsf{I}\setminus\mathsf{G}}\hat{\zeta}^2_*}}\right\} = \alpha, \tag{20}$$

if we let

$$c_\alpha = {\chi^2}_{1-\alpha}\left(k-k_g, \frac{n\Delta_{\mathsf{I}\setminus\mathsf{G}}\zeta^2_*}{4}\right) - \frac{n\Delta_{\mathsf{I}\setminus\mathsf{G}}\zeta^2_*}{4}.$$

To perform feasible inference one will need to estimate $\Delta_{\mathsf{I}\setminus\mathsf{G}}\,\zeta^2_*$. Again this will be via the connection to a non-central F-distribution, Equation 19. One can either find an upper quantile directly, or use a KRS-type estimator, which for the hedged case are

$$\begin{aligned}\delta_0 &= \frac{(n-k-2)}{n-1}\Delta_{\mathsf{I}\setminus\mathsf{G}}\hat{\zeta}^2_* - \frac{k-k_g}{n}\left(1+\frac{n}{n-1}\hat{\zeta}^2_{*,\mathsf{G}}\right),\\ \delta_1 &= max\left(\delta_0, 0\right),\\ \delta_2 &= max\left(\delta_0, \frac{2}{k-k_g+2}\frac{n-k-2}{n-1}\Delta_{\mathsf{I}\setminus\mathsf{G}}\hat{\zeta}^2_*\right).\end{aligned} \tag{21}$$

[1] *n.b.* the standard definition of non-centrality parameter in the time Graybill and Marsaglia wrote their paper is different from the one we use today by a factor of 1/2.

Checking Coverage As in the unhedged case, we first perform simulations where the population parameter $\Delta_{\mathsf{I}\setminus\mathsf{G}}\ \zeta_*^2$ is known, to assess the effects of Approximation 8. In our simulations, we set $k_g = k/2$, and let G be the first k_g rows of the identity matrix. We set $\boldsymbol{\mu} = c\mathbf{1}$ and $\Sigma = \mathsf{I}$. We perform $100,000$ simulations for different values of $\Delta_{\mathsf{I}\setminus\mathsf{G}}\zeta_*^2$, k and n, computing ζ_a for the hedged portfolio, as well as $\Delta_{\mathsf{I}\setminus\mathsf{G}}\ \hat{\zeta}_*^2$. We compute the lower 0.05 bound using knowledge of $\Delta_{\mathsf{I}\setminus\mathsf{G}}\ \zeta_*^2$ and compute the empirical type I rate over the $100,000$ simulations, which we plot versus n/k in Figure 3. Again we see that the nominal type I rate is nearly achieved when $n > 100k$ or so.

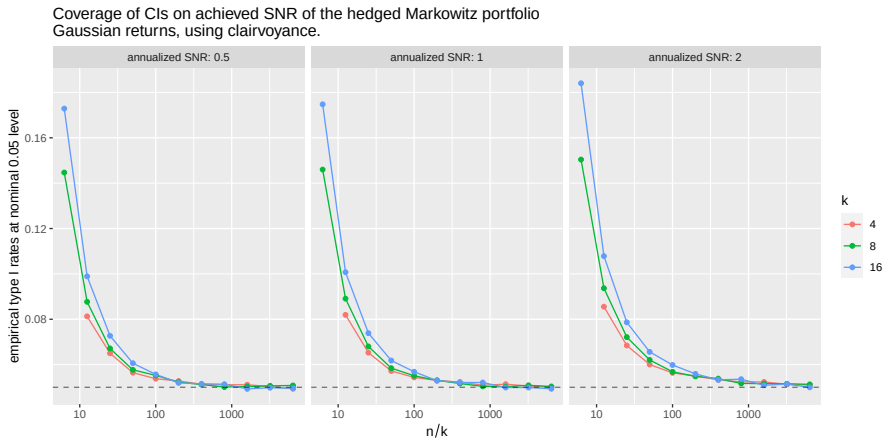

Figure 3: The empirical type I rate, over 100,000 simulations, of a one-sided confidence bound for ζ_a are shown more tk.

3 Examples

Fama French 4 Factor Returns We consider a portfolio constructed on the 'Market', size (SMB), value (HML) and momentum (HMD) portfolios described by Fama and French, *inter alia*, with data compiled and published by Kenneth French. [4, 3, 5] The set consists of $n = 1104\,\text{mo.}$ of data, from Jan 1927 through Dec 2018. We observe $\hat{\zeta}_*^2 = 0.098\,\text{mo.}^{-1}$. From this we compute $\delta_2 = 0.094\,\text{mo.}^{-1}$. Plugging in $\sqrt{\delta_2}$ for ζ_* we compute a two-sided 95% confidence bound on ζ_a as $[0.234, 0.353]\ \text{mo.}^{-1/2}$. By comparison, via the connection to the F distribution, we compute 95% confidence intervals on ζ_* as $[0.247, 0.369]\ \text{mo.}^{-1/2}$.

4 Discussion

Testing indicates the confidence bound exhibits closer to nominal coverage than the known asymptotic bounds for reasonable n and k. Further work should naturally focus on mitigating the effects of the approximation $\hat{\Sigma} \approx \Sigma$, and finding a coverage guarantee of the plug-in estimator. We also anticipate that this confidence bound procedure can be adapted to deal with hedging constraints and conditional expectation models.

References

[1] T. W. Anderson. *An Introduction to Multivariate Statistical Analysis*. Wiley Series in Probability and Statistics. Wiley, 2003. ISBN 9780471360919. URL http://books.google.com/books?id=Cmm9QgAACAAJ.

[2] Mengmeng Ao, Yingying Li, and Xinghua Zheng. Solving the Markowitz optimization problem for large portfolios. URL http://fmaconferences.org/SanDiego/Papers/MAXSER_MengmengAo.pdf.

[3] Mark M Carhart. On persistence in mutual fund performance. *Journal of Finance*, 52(1):57, 1997. URL http://www.jstor.org/stable/2329556.

[4] Eugene F. Fama and Kenneth R. French. The cross-section of expected stock returns. *Journal of Finance*, 47(2):427, 1992. URL http://www.jstor.org/stable/2329112.

[5] Kenneth French. Data library. Privately Published, 2017. URL http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.

[6] Narayan C. Giri. On the likelihood ratio test of a normal multivariate testing problem. *The Annals of Mathematical Statistics*, 35(1):181–189, 1964. doi: 10.1214/aoms/1177703740. URL http://projecteuclid.org/euclid.aoms/1177703740.

[7] Franklin A. Graybill and George Marsaglia. Idempotent matrices and quadratic forms in the general linear hypothesis. *The Annals of Mathematical Statistics*, 28(3):678–686, 1957. ISSN 00034851. URL http://www.jstor.org/stable/2237227.

[8] Tatsuya Kubokawa, C. P. Robert, and A. K. Md. E. Saleh. Estimation of noncentrality parameters. *Canadian Journal of Statistics*, 21(1):45–57, Mar 1993. URL http://www.jstor.org/stable/3315657.

[9] Richard O. Michaud. The Markowitz optimization enigma: is 'optimized' optimal? *Financial Analysts Journal*, pages 31–42, 1989. URL http://newfrontieradvisors.com/Research/Articles/documents/markowitz-optimization-enigma-010189.pdf.

[10] Dirk Paulsen and Jakob Söhl. Noise fit, estimation error and a Sharpe information criterion. *Quantitative Finance*, 0(0):1–17, 2020. doi: 10.1080/14697688.2020.1718746. URL https://doi.org/10.1080/14697688.2020.1718746.

[11] Steven E. Pav. Asymptotic distribution of the Markowitz portfolio. Privately Published, 2013. URL http://arxiv.org/abs/1312.0557.

[12] Steven E. Pav. Bounds on portfolio quality. Privately Published, 2014. URL http://arxiv.org/abs/1409.5936.

[13] Steven E. Pav. Untitled book. Forthcoming, 2020.