Bayesian Estimation for Sharpe Ratios under Selection Bias

Joseph Mulligan
Imperial College London
Qube Research & Technologies

QMF 2024

Overview

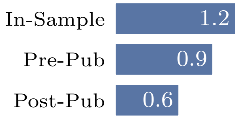
Question: How do we estimate a Sharpe ratio, accounting for selection bias?

Proposition: Derive the posterior distribution for the Sharpe ratio using a bias-adjusted likelihood, jointly fitting prior hyperparameters to observed in-sample and out-of-sample Sharpe ratios.

Findings: This particularly improves estimates for short backtests, and gives non-linear adjustments which diminish for high Sharpe ratios.

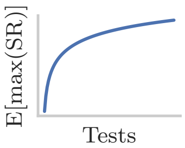
Motivation

- Out-of-sample performance is often worse than in-sample performance. Why?



Dataset: 206 predictors from [CZ22].

- One explanation: Multiple testing inflates statistics, but tracking the number of tests (as required) is rarely done [BL14; HLZ16].



Expected Maximum Sharpe after N tests [BL14].

Instead, we model *why* researchers conduct multiple tests:
To search for higher Sharpe ratios.

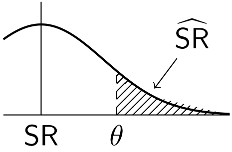
Chen and Zimmermann [CZ20]:

- Adjusts in-sample return for **publication bias**
- Focuses on estimation **without** out-of-sample data
- Finds **small publication bias** relative to post-pub decay

Our Contribution:

- Seek the **best Bayesian estimate** for **out-of-sample Sharpe ratios**
- Jointly fit hyperparameters to IS **and** OOS SRs
- Utilise a **conditional likelihood** for the observed sample statistics

Imagine we have selected a strategy based on a sample Sharpe ratio $\widehat{\text{SR}}$ which clears a minimum threshold θ.



The sample Sharpe ratio $\widehat{\text{SR}}$ has been **biased upwards** from the true value SR due to selection from a **truncated distribution**.

- The sample Sharpe ratio is a scaled t-statistic, $t = \sqrt{T}\widehat{\mathsf{SR}}$.
- Assuming the payoffs of the strategy are Normal, the sample Sharpe ratio has distribution $\sqrt{T}\widehat{\mathsf{SR}} \sim t_{T-1}\left(\sqrt{T}\mathsf{SR}\right)$.

The truncated distribution of the biased sample Sharpe ratio is given by

$$f_{\widehat{\mathsf{SR}}|\widehat{\mathsf{SR}}>\theta}\left(\widehat{\mathsf{SR}} \mid \mathsf{SR}, T, \theta\right) = \frac{f_{\widehat{\mathsf{SR}}}\left(\widehat{\mathsf{SR}} \mid \mathsf{SR}, T\right) \mathbb{1}_{\widehat{\mathsf{SR}}>\theta}}{1 - F_{\widehat{\mathsf{SR}}}\left(\theta \mid \mathsf{SR}, T\right)},$$

where $f_{\widehat{\mathsf{SR}}}$ and $F_{\widehat{\mathsf{SR}}}$ denote the original density and CDF of the sample Sharpe ratio.

Proposition

Let SR $= 0$*, then the expected sample Sharpe is given by,*

$$\mathbb{E}\left[\widehat{\mathsf{SR}} \;\middle|\; \widehat{\mathsf{SR}} > \theta, \mathsf{SR} = 0\right] = \frac{U}{\sqrt{T}}$$

where

$$U = \kappa \frac{T-1}{T-2}\left(1 + \frac{T}{T-1}\theta^2\right)^{-\frac{T-2}{2}}$$

$$\kappa = \frac{\Gamma\left(\frac{T}{2}\right)}{\alpha_0 \Gamma\left(\frac{T-1}{2}\right)\sqrt{(T-1)\pi}}, \qquad \alpha_0 = 1 - F_t\left(\sqrt{T}\theta; T-1\right).$$

Expected Bias Over True Sharpe Ratio

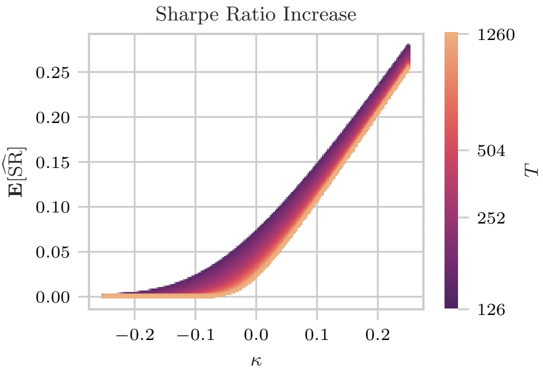
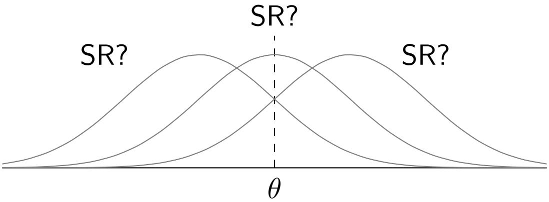


Figure: Sharpe ratio inflation by θ and T. Values are in terms of daily Sharpe ratios, and T is in days.

But what is the truth?

In reality, the true SR is unknown, and we don't know if our threshold is above or below the truth.



So how could we estimate the true Sharpe, given an observed sample Sharpe ratio and a threshold?

We need:

- A prior for the Sharpe ratio $p\left(\mathsf{SR} \mid \Theta\right)$,
- An appropriate likelihood which is aware of the threshold bias $p\left(\widehat{\mathsf{SR}} \mid \mathsf{SR}, \widehat{T}, \Theta, \mathsf{acc}\right)$.

We can then compute the posterior for the true Sharpe ratio,

$$p(\mathsf{SR} \mid \widehat{\mathsf{SR}}, \widehat{T}, \Theta, \mathsf{acc}) \propto p\left(\mathsf{SR} \mid \Theta\right) p\left(\widehat{\mathsf{SR}} \mid \mathsf{SR}, \widehat{T}, \Theta, \mathsf{acc}\right).$$

If the payoffs are Normal, then a Normal-Inverse-Gamma prior is the natural choice for the variance and the Sharpe ratio, [Pav21]

$$\sigma^2 \sim \Gamma^{-1}\left(\frac{m_0}{2}, \sigma_0^2 \frac{m_0}{2}\right),$$

$$\mathsf{SR} \mid \sigma^2 \sim \mathcal{N}\left(\frac{\mu_0}{\sigma}, \frac{1}{n_0}\right),$$

which yields a marginal prior for the Sharpe ratio,

$$\sqrt{n_0}\mathsf{SR} \sim \lambda'\left(\sqrt{n_0}\mathsf{SR}_0, m_0\right),$$

where λ' denotes Lecoultre's lambda prime distribution.

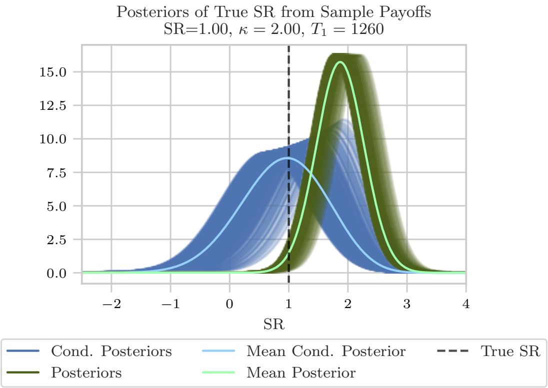
As in [CZ20] we extend to a soft rather than hard threshold,

$$p_{\mathsf{acc}}\left(x \mid \kappa, \ell\right) = \frac{1}{1 + \exp(-\ell(x-\kappa))},$$

and we already know the sample distribution of the Sharpe ratio, $p_{\widehat{\mathsf{SR}}}(\widehat{\mathsf{SR}} \mid \mathsf{SR}, \widehat{T})$, so our likelihood is given by

$$p\left(\widehat{\mathsf{SR}} \mid \mathsf{SR}, \widehat{T}, \Theta, \mathsf{acc}\right) = \frac{p_{\widehat{\mathsf{SR}}}(\widehat{\mathsf{SR}} \mid \mathsf{SR}, \widehat{T}) p_{\mathsf{acc}}(\widehat{\mathsf{SR}} \mid \kappa, \ell)}{\mathbb{E}_{\widehat{\mathsf{SR}}}\left[p_{\mathsf{acc}}(\widehat{\mathsf{SR}} \mid \kappa, \ell)\right]}.$$

We can therefore use a sample Sharpe ratio and a threshold to recover the posterior for the true Sharpe ratio:



However, a problem remains: **How do we pick Θ?**

We propose using empirical Bayes to estimate the hyperparameters. This has a few key advantages:

1. We "leverage" a dataset of observed in-sample *and* out-of-sample Sharpe ratios to fit our parameters.
2. This enables us to correct for the selection bias, and also improve performance estimates in a limited information environment.
3. This is particularly useful for short backtests, where we have limited information to accurately estimate the Sharpe ratio of the strategy.

- We can use maximum likelihood estimation to estimate the parameters in $\Theta = \{n_0, m_0, \mu_0, \sigma_0, \kappa, \ell\}$ from a dataset of strategies.
- If we have a dataset of strategies, we likely have both in-sample and out-of-sample performance for each strategy.
- We would like to use all the data available to us, so we need the joint density of the in-sample and out-of-sample Sharpe ratio for a singular strategy:

$$p\left(\widehat{\mathsf{SR}}_i, \widetilde{\mathsf{SR}}_i \mid \widehat{T}_i, \widetilde{T}_i, \Theta, \mathsf{acc}\right)$$

To get the joint density of $\widehat{\mathsf{SR}}, \widetilde{\mathsf{SR}}$ we use the joint density of $\widehat{\mu}, \widehat{\sigma}^2, \widetilde{\mu}, \widetilde{\sigma}^2$ and apply the necessary transform.

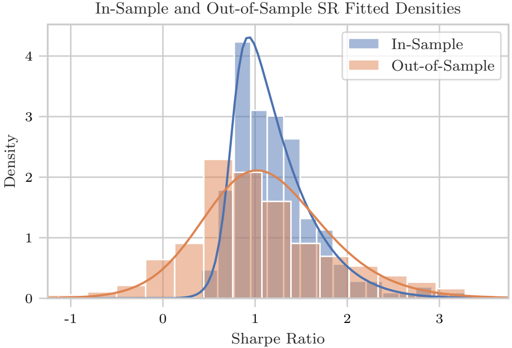
The joint density of the latter is given by,

$$p\left(\widehat{\mu}, \widehat{\sigma}^2, \widetilde{\mu}, \widetilde{\sigma}^2 \mid \widehat{T}, \widetilde{T}, \Theta, \mathsf{acc}\right) = q(\widetilde{\mu}, \widetilde{\sigma}^2 \mid \widetilde{T}, \Theta_1) \frac{q(\widehat{\mu}, \widehat{\sigma}^2 \mid \widehat{T}, \Theta_0) p_{\mathsf{acc}}(\widehat{\mu}/\widehat{\sigma} \mid \kappa, \ell)}{\int_0^\infty \int_{-\infty}^\infty q(m, s^2 \mid \widehat{T}, \Theta_0) p_{\mathsf{acc}}(m/s \mid \kappa, \ell) \, \mathrm{d}m \, \mathrm{d}s^2},$$

where $q(x, y)$ is the joint likelihood of the sample mean and variance given a Normal-Inverse-Gamma prior, which we have marginalised out, and Θ_j refers to either the prior, or posterior updated, parameters of the prior.

Voila Encore!

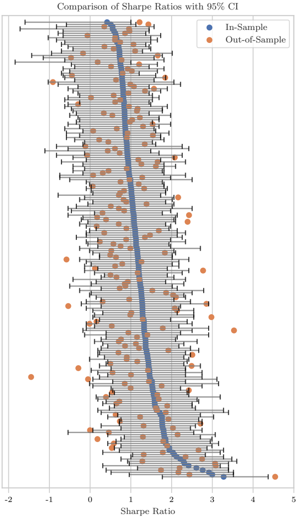
We can now fit our parameters to some dataset...



Dataset: 206 predictors from [CZ22].

Voila Encore!

...and use this to compute confidence intervals for the out-of-sample performance.



Dataset: 206 predictors from [CZ22].

1. We can fit a Bayesian model to a dataset of in-sample and out-of-sample Sharpe ratios.
2. The model takes into account the selection criteria used to choose strategies for trading.
3. And we can use this model to improve estimates of out-of-sample performance of new candidate strategies!

Bibliography I

[BL14] David H. Bailey and Marcos López De Prado. "The Deflated Sharpe Ratio: Correcting for Selection Bias, Backtest Overfitting, and Non-Normality". In: *The Journal of Portfolio Management* 40.5 (Sept. 2014), pp. 94–107. ISSN: 0095-4918, 2168-8656. DOI: 10.3905/jpm.2014.40.5.094.

[CZ20] Andrew Y Chen and Tom Zimmermann. "Publication Bias and the Cross-Section of Stock Returns". In: *The Review of Asset Pricing Studies* 10.2 (June 2020). Ed. by Jeffrey Pontiff, pp. 249–289. ISSN: 2045-9920, 2045-9939. DOI: 10.1093/rapstu/raz011.

[CZ22] Andrew Y. Chen and Tom Zimmermann. "Open Source Cross-Sectional Asset Pricing". In: *Critical Finance Review* 11.2 (May 2022), pp. 207–264. ISSN: 2164-5744, 2164-5760. DOI: 10.1561/104.00000112.

[HLZ16] Campbell R. Harvey, Yan Liu, and Heqing Zhu. ". . . and the Cross-Section of Expected Returns". In: *The Review of Financial Studies* 29.1 (Jan. 2016), pp. 5–68. ISSN: 0893-9454. DOI: 10.1093/rfs/hhv059.

[Pav21] Steven E. Pav. *The Sharpe Ratio: Statistics and Applications*. New York: Chapman and Hall/CRC, Sept. 2021. ISBN: 978-1-00-318105-7. DOI: 10.1201/9781003181057.