Uncertainty of Returns in Light of Forward Looking Priors

Gabriel Stechschulte; gabriel.stechschulte@stud.hslu.ch

Lecturers

Thomas Ankenbrand

Denis Bieri

Time Series Analysis in Finance W.MSCIDS_DE_TSA01.F2101

Master of Science, Applied Information and Data Science

Expected Returns in Light of Forward Looking Priors submitted as part of the requirements for the MSc. in Applied Information and Data Science at the school of business, Lucerne University of Applied Sciences and Arts

05.28.2021

Table of Contents

Introduction

In finance and econometrics, most practitioners leverage common statistical methods to analyze various financial instruments' return characteristics. Namely, "*are the returns desirable for the risk I am taking on*?". Statistical methods to analyze such a question include the mean, standard deviation, skewness and kurtosis, Sharpe ratio, and drawdown. Advanced econometric models such as the autoregressive integrated moving average (ARIMA), generalized autoregressive conditional heteroskedasticity (GARCH), vector autoregression (VAR), and various forms of multiple and generalized linear regression are also used to form hypotheses of risk, return characteristics, and whether or not to invest in certain financial products.

With these classical approaches, the parameter(s) of interest, μ (monthly return) and σ (monthly volatility), are treated as unknown but fixed. The parameter(s) being fixed means when it comes to hypothesis testing under classical statistical methods, probabilities are not returned. Rather, the summary output is that the null hypothesis fails to be rejected or is rejected. As well, in classical statistics, is it often the case that the errors of models are to be normally distributed; an assumption in practice that is usually difficult to make valid. Bayesian inference, which treats the parameters μ and σ as unknown but as random variables, offers an advantage over classical statistical methods as 1.) the ability to inject a prior belief and 2.) a posterior distribution is returned as output (Bloom & Orloff, 2018). This has numerous benefits as you can quantify uncertainty in the form of probabilities with credible intervals and you can leverage monte carlo markov chain (MCMC) sampling since distributions are returned as the output (posterior).

Literature and Introduction to Bayesian Statistics

Applications of bayesian statistics and modeling are being applied in a wide range of fields such as optimal monetary policy (Cogley et al., 2011), financial economics (Cremers, 2002), deep learning (Wilson & Izmailov, 2020), and even in the search for missing Malaysia flight MH370 (Davey et al., 2016). The rise of Bayesian inference can be partly attributed to the power of computation increasing in the past two decades as well as a decrease in price for more powerful and modern central processing units (CPUs) and graphical processing units (GPUs). Also, well documented open source packages such as JAGS and RJAGS in the R environment and pyMC3 in the Python environment have been developed primarily for bayesian analysis and MCMC sampling - democratizing permission technologies.

As stated above in the introduction, there are two schools of statistics; 1.) Frequentist / Classical and 2.) Bayesian. The Bayesian school comes from Bayes theorem, published in 1763, which links the prior and conditional probability to the posterior probability:

$$Pr(Y|X) = \frac{Pr(X|Y)Pr(Y)}{Pr(X)}$$

Which can be extended to continuous variables (parameters of interest in this example) with their respective probability density functions (pdf) and the observational data:

$$f(\theta|Y) = \frac{f(Y|\theta)f(\theta)}{f(Y)}$$

Where $f(\theta|Y)$ is the posterior distribution, $f(Y|\theta)$ is the likelihood function, $f(\theta)$ is the prior, and $f(Y)$ is the evidence (marginal likelihood). The posterior is the updated conditional distribution after incorporating the prior *and* seeing the data. The likelihood describes how θ and Y are related through the conditional. The prior distribution is saying, "we might know something about the possible values of parameter(s) of interest *prior* to seeing the data". Incorporating expert knowledge to guide the parameter's prior probabilities, which can improve the posterior distribution estimations. Lastly, the marginal likelihood, the denominator, is just a normalizing constant, and thus can be ignored since the posterior is proportional to the product of the likelihood and the prior:

$$f(\theta|Y) \propto f(Y|\theta)f(\theta)$$

Where θ is the parameters of interest, μ and σ, and Y is the observed data (monthly returns from a given time period).

The models above allow Bayesian's to model uncertainty and interpret probability as a measure of *believability* or *confidence* in an event occuring. I.e., There is a 75% chance you have a deadly illness. They are able to do this since the parameters are unknown but treated as random variables. The beautiful idea behind Bayesian statistics is the rich intuitiveness behind it. John Maynard Keynes said it best, "When the facts change, I change my mind." (A Quote by John Maynard Keynes, n.d.). Distributions are convenient methods to express current beliefs (prior) and the improved state of belief (posterior) - and it just so happens these are both encapsulated in distributions. Thus, we are able to leverage Markov Chain Monte Carlo (MCMC) to sample thousands of data points from the posterior to help us get better inference on the uncertainty of our parameters.

Methodology

In this analysis, an exchange traded fund (ETF) returns downloaded from the Yahoo Finance API operating in the technology sector will be modeled using JAGS and RJAGS, programs for performing bayesian analysis and inference with MCMC. After loading and preprocessing the data in R, a normal likelihood uninformative prior will be modeled using the returns data from 2015-01-1 until 2020-02-01 to ensure the model "learns" the parameters of interest. Then, a normal likelihood informative prior (see table 1 in appendix) will be modeled using the returns data from 2018-01-01 until 2020-02-01 with a hypothetical expert injecting her beliefs as the prior for the model. This prior will reflect the expert's knowledge of the Coronavirus pre-market crash and the time frame is adjusted to give more weight to the prior belief (distribution) and less reliance on past data. Model diagnostics such as investigating the MCMC trace plots, plotting correlation functions to ensure there is no evident correlation in the chain as a function of the lag (i.e., If the ACF takes too long to decay to 0, the chain exhibits a high degree of dependence and will tend to get stuck in place) will be completed for each model.

Although she has been experiencing positive returns over the past months, she expects the Coronavirus to have a negative impact on the technology sector. Thus, the hypothesis is that she expects the probability of mean monthly continuous returns (μ >= 0.05%) to decrease and the probability of mean monthly volatility (σ > 2%) to increase for the ETF in the next five months ahead. As a result of a posterior distribution being returned, MCMC sampling will be used to gather 1,000 samples from four chains to perform inference and probability calculations. Subsequently, we are interested in the uncertainty of the returns and volatility of the mutual fund before and in light of our beliefs being incorporated into the model. By incorporating a forward looking prior, we can compute the uncertainty of expected returns of our current investment in SPY. Continuous returns can be calculated by:

$$r_t = ln(\frac{P_t}{P_{t-1}})$$

Table 1: Modeling Workflow

Step	Uninformative Model	Informative Model
1.	Retrieve SPY prices (2015 - 2020-02-01) from API and compute continuous returns	Retrieve SPY prices (2018 - 2020-02-01) from API and compute continuous returns
2.	Model parameters μ, σ as random variables and specify uninformative prior	Model parameters μ, σ as random variables and specify informative prior
3.	Run JAGS model to update posterior distribution	Run JAGS model to update posterior distribution
4.	Use MCMC to simulate 1000 samples of parameters from the posterior distribution	Use MCMC to simulate 1000 samples of parameters from the posterior distribution
5.	Perform model diagnostics and summary statistics	Perform model diagnostics and summary statistics
6.	Perform inference of expected returns and probabilities of mean monthly return and mean monthly volatility	Perform inference of expected returns and probabilities of mean monthly return and mean monthly volatility

Conclusion

The goal of the uninformative prior is to develop a "baseline" model such that it would effectively learn the parameters μ and σ that generated the observational data. Subsequently, given the appropriate model, the probability of experiencing mean monthly returns (Pr(μ > 0.05%)) and reasonable risk (Pr(σ < 2.0%)) for the time period was calculated. Upon establishing the baseline model, an informative model with the experts prior belief was modeled (figure 1 of appendix). As the informative model contains the majority of the valuable information and where inference takes place, the model diagnostics can be found in figure 1, 2, 3, and 4 in the appendix. Below, in table 2, are the results and inference from our modeling workflow.

Table 2: Bayesian Uninformative and Informative Model results

	Uninformative Model	Informative Model
Time Period	2015-01-01 to 2020-02-01	2018-01-01 to 2020-02-01
Posterior mean μ *95% Credible Interval*	0.0046 [-0.003, 0.012]	-0.003 [-0.014, 0.008]
Realized mean monthly return μ	0.0096	0.011
Pr(μ >= 0.05%)	0.494	0.395
Pr(μ < 0.0%)	0.439	0.539
Posterior mean σ *95% Credible Interval*	0.029 [0.026, 0.034]	0.0296 [0.024, 0.036]
Realized mean monthly σ	0.035	0.111
Pr(σ >= 2.0%)	0.627	0.627

Note. Realized mean and volatility for informative prior was calculated for the time period 2020-02-01 to 2020-07-01. This is done to see if the realized calculations fall within the credible intervals which in turn gives us an understanding if our subjective priors are in line with the forward looking data

As the results show, by incorporating forward looking prior beliefs, the probability of achieving monthly returns greater than or equal to 0.05% decreased from ~50% to ~40%. Subsequently, the probability of enduring more volatility, remarkably, stayed the same at ~62% (volatility for each time period was high as can be seen by the realized σ). Although the expert expected lower returns and higher volatility, inferring the realized mean returns and volatility from the time period 2020-02-01 to 2020-07-01 to reflect the forward looking time frame with the credible intervals of the informative model from the time frame 2018-01-01 to 2020-02-01, tell us the realized values of μ and σ were above the upper range of the 95% credible interval. We can conclude that the realized returns and volatility were higher than the expert expected and that her prior belief did not capture this.

Due to these findings, we hypothesize that the expert may have been too risk averse and or biased to believe the long term negative impact of Covid on the equity markets, and in particular, the SPY ETF. Particularly, the investor may have underestimated the magnitude of fiscal and monetary policy, the Cantillon effect, and investor sentiment. A forward looking prior, although subjective, can help investors understand the probability of risk and returns of particular asset classes if they choose the prior *wisely*.

We have discussed the various benefits of utilizing Bayesian statistics over frequentist statistics in analyzing future equity returns. As a result, we modeled the returns of a technology ETF using an uninformative prior and an informative forward looking prior. With the benefit of hindsight, the prior reflected the expert's knowledge of the Covid-crisis negatively impacting the markets. As a result of this forward looking prior, we used MCMC to generate 1,000 samples from our posterior distribution of θ to model the uncertainty of our model's parameters μ and σ. Nonetheless, the returns expected in the future decreased with the probability of experiencing negative mean monthly returns representing the majority (> 50%) which can be visualized in figure 5 of the appendix. This insight could be used to allow the investor to sell/open other positions, enter short positions and or engage in risk management techniques.

Discussion

The use of subjective priors in Bayesian inference is widely criticized due to reproducibility - i.e., There is no one single method in determining the prior. Therefore, depending on the human, different outcomes may be achieved. But, when the prior is chosen diligently, this is where the beauty of Bayesian inference comes into play due to its impact on the posterior distribution. Likewise, more formal rules may be used in determining the prior outlined by Kass & Wasserman, 1996. As well, there is also empirical Bayesian analysis which uses other statistical methods and inference to hypothesize a more objective prior. Such methods could perhaps be used to identify market factors such as interest rates, central bank activity, monetary and fiscal policy, investor sentiment, etc. that influence the returns of a particular sector or asset class. Incorporating these other methods into a sequence of models could be used to define a more objective prior.

Bibliography

A quote by John Maynard Keynes. (n.d.). Retrieved May 7, 2021, from https://www.goodreads.com/quotes/353440-when-the-facts-change-i-change-my-mind---what

Cogley, T., De Paoli, B., Matthes, C., Nikolov, K., & Yates, T. (2011). A Bayesian approach to optimal monetary policy with parameter and model uncertainty. *Journal of Economic Dynamics and Control*, 35(12), 2186–2212. https://doi.org/10.1016/j.jedc.2011.02.006

Cremers, K. J. M. (2002). Stock Return Predictability: A Bayesian Model Selection Perspective. *The Review of Financial Studies*, 15(4), 1223–1249.

Davey, S., Gordon, N., Holland, I., Rutten, M., & Williams, J. (2016). *Bayesian Methods in the Search for MH370*. Springer Singapore. https://doi.org/10.1007/978-981-10-0379-0

Kass, R. E., & Wasserman, L. (1996). The Selection of Prior Distributions by Formal Rules. *Journal of the American Statistical Association*, 91(435), 1343–1370. https://doi.org/10.2307/2291752

Bloom, J., Orloff J. (2018). Comparison of Frequentist and Bayesian Inference. *MIT18_05S14_Reading*. Retrieved May 7, 2021, from https://ocw.mit.edu/courses/mathematics/18-05-introduction-to-probability-and-statistics-spring-2014/readings/MIT18_05S14_Reading20.pdf

Wilson, A. G., & Izmailov, P. (2020). Bayesian Deep Learning and a Probabilistic Perspective of Generalization. *ArXiv:2002.08791* [Cs, Stat]. http://arxiv.org/abs/2002.08791

Appendix

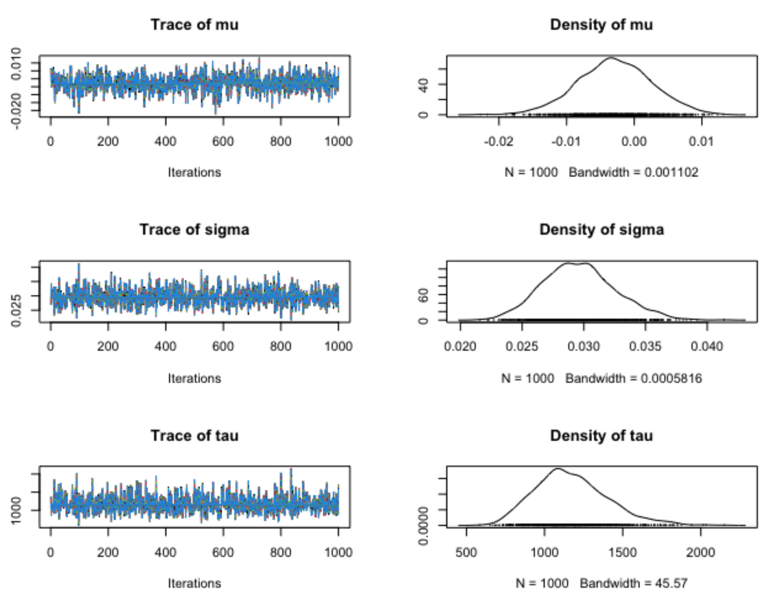
Figure 1

Prior Distribution

Note. Simulated prior distribution for τ and μ

Figure 2

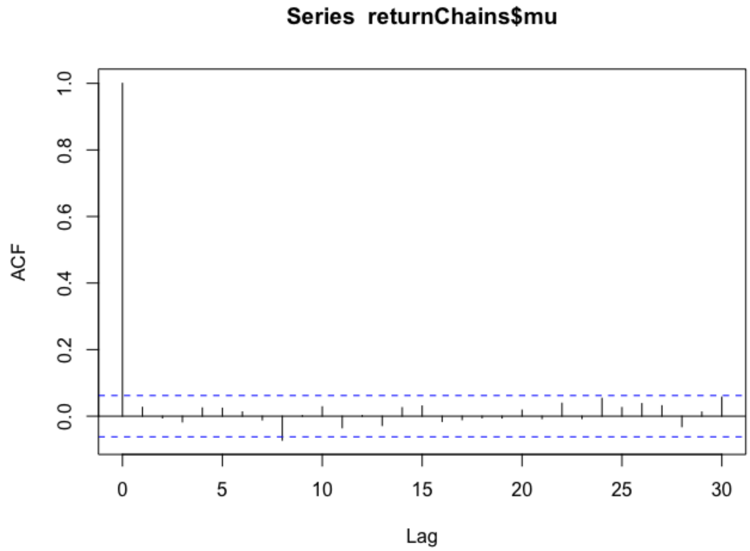
Posterior MCMC trace plots



Note. Posterior MCMC trace plots for informative model resemble a white noise process

Figure 3

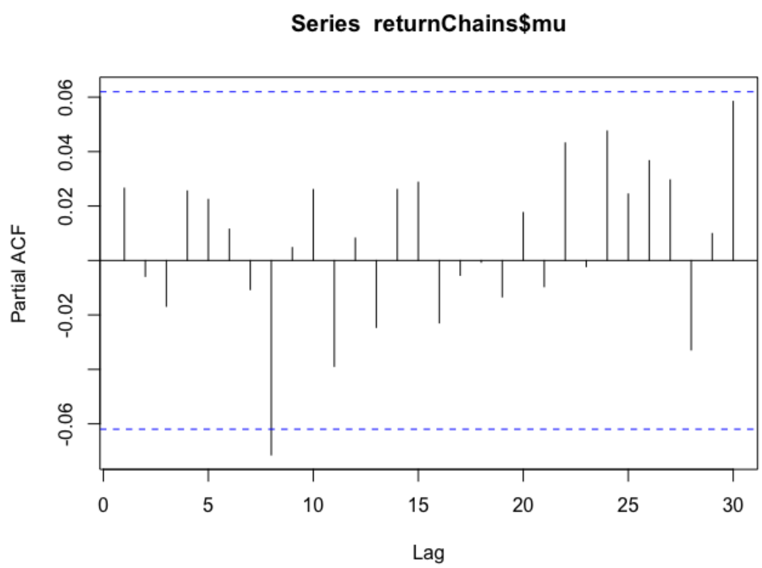
Autocorrelation Function (ACF) of the mu chain



Note. ACF shows there is no direct and or indirect dependence of the values in mu for the chain

Figure 4

Partial Autocorrelation function (PACF) of the mu chain



Note. PACF indicates a weakly significant lag at ~7. This has been ignored when performing inference as the overall PACF process does not exhibit any clear AR processes.

Figure 5

Posterior Predictive Distribution for mean monthly returns