Statistics and Probability Letters 134 (2018) 106–113

Contents lists available at ScienceDirect

Statistics and Probability Letters

journal homepage: www.elsevier.com/locate/stapro

# ‘Purposely misspecified’ posterior inference on the volatility of a jump diffusion process

Ryan Martin [a], Cheng Ouyang [b,*], Francois Domagni [b]

[a] *Department of Statistics, North Carolina State University, United States*
[b] *Department of Mathematics, Statistics, and Computer Science, University of Illinois at Chicago, United States*

A R T I C L E  I N F O

*Article history:*
Received 10 March 2017
Received in revised form 7 September 2017
Accepted 18 October 2017
Available online 11 November 2017

*Keywords:*
Bernstein–von Mises theorem
Cramér–Rao lower bound
Credible interval
Gibbs posterior
Uncertainty quantification

A B S T R A C T

Bayesian analysis requires prior distributions for all model parameters, whether of interest or not. This can be a burden, for a number of reasons, especially when the nuisance parameters are high- or infinite-dimensional, so there is motivation to find a way around this without completely abandoning the Bayesian approach. Here we consider a general strategy of working with a *purposely misspecified model* to avoid dealing directly with nuisance parameters. We focus this investigation on an interesting and challenging problem of inference on the volatility of a jump diffusion process based on discrete observations. If we simply ignore the jumps, we can work out precisely the asymptotic behavior of the Bayesian posterior distribution based on the misspecified model. This result suggests some simple adjustments to correct for the effects of misspecification, and we demonstrate that a suitably corrected version of our purposely misspecified posterior leads to inference on the volatility that is asymptotically optimal.



## 1. Introduction

The Bayesian approach, i.e., where a prior distribution for the unknown parameters is updated to a corresponding posterior based on observed data and Bayes’ formula, provides a powerful framework for statistical inference. An essential feature of a Bayesian analysis is that it requires a probability model for all uncertain or variable quantities, namely, prior distributions for all unknown parameters and a likelihood function corresponding to all observables. Often some of the parameters are nuisance, i.e., needed to specify a suitable likelihood but not of direct inferential interest, so it is tempting to avoid specifying priors and computing posteriors for these, to focus resources on the interest parameters only. This is especially true in so-called semiparametric problems where primary interest is in some finite-dimensional parameter, but the model itself involves a complex infinite-dimensional nuisance parameter. The goal of this paper is to describe, through a particular semiparametric example in finance, a new approach by which a Bayesian analysis can directly target the interest parameters by considering a *purposely misspecified* model. The idea is that certain biases will appear as a result of misspecification but, if these biases can be characterized, then some simple corrections to the posterior can be made and, therefore, we avoid both dealing directly with complex nuisance parameters and the negative consequences of misspecification.

The particular setting in which we will carry out our “misspecification on purpose” investigation is in the modeling of asset prices over time. Compared to the classical Black–Scholes models (e.g. Musiela and Rutkowski, 2005), based solely on a continuous Brownian motion, jump diffusion models – which have both continuous and jump parts – have received

* Corresponding author.
*E-mail addresses:* rgmarti3@ncsu.edu (R. Martin), couyang@uic.edu (C. Ouyang), fdomag2@uic.edu (F. Domagni).

https://doi.org/10.1016/j.spl.2017.10.013
0167-7152/

considerable attention in the last two decades, largely because they can accommodate the rapid, seemingly discontinuous changes in asset prices often observed in applications. Having the two parts is important because, as several authors have concluded, neither a purely-continuous nor purely-jump model is sufficient for real applications (e.g., Aït-Sahalia and Jacod, 2009, 2010; Barndorff-Nielsen and Shepard, 2006; Podolskij, 2006). Indeed, by comparing the observed behavior of at-the-money and out-of-the-money call option prices near expiration with their analogous theoretical behavior, Carr and Wu (2003) and Medvedev and Scaillet (2007) argued that both continuous and jump components are necessary to explain the implied volatility behavior of S&P500 index options. In this paper, we consider a continuous-time process $X = (X_t : t \in [0, T])$ over a fixed and finite time horizon $[0, T]$ that can be decomposed as

$$X_t = \beta t + \theta^{1/2} W_t + J_t, \quad t \in [0, T], \tag{1}$$

where $\beta t + \theta^{1/2} W_t$ is a continuous diffusion – with $\beta$ and $\theta$ the drift and volatility coefficients, respectively, and $(W_t : t \in [0, T])$ a standard Brownian motion – and $J = (J_t : t \in [0, T])$ is a pure jump process with finite jump activity, independent of $W$. If the entire process $X$ were observable, then we could immediately identify and extract the jumps, reducing the problem to a relatively simple one where the continuous part and jump parts are analyzed separately. However, in our application, we only observe $X$ at $n$ fixed times $0 < t_1 < t_2 < \cdots < t_n < T$, like in, e.g., Aït-Sahalia and Jacod (2009) and Figueroa-López (2009). Having only discrete-time observations means that the continuous and jump parts cannot be disentangled with certainty, so a separate analysis of the continuous and jump parts is not possible.

The model (1) is characterized by $\beta, \theta$, and whatever parameters are needed to describe the distribution of $J$. Certainly all of these parameters could be of interest, but here we take the volatility coefficient, $\theta$, a fundamentally important measure of uncertainty or risk (Musiela and Rutkowski, 2005), as our parameter of interest; this makes both the drift and the distribution of $J$ nuisance parameters. Frequentist approaches are available for inference on $\theta$, e.g., in Aït-Sahalia and Jacod (2014), but, as mentioned above, we will be pursuing a Bayesian approach here.

A proper Bayesian approach to this problem requires a prior distribution for both the interest and the nuisance parameters; the data analyst will then get a full joint posterior distribution and proceed to integrate out the nuisance parameters to get a marginal posterior distribution to be used for inference on $\theta$. Specifying a prior distribution for $(\beta, \theta)$ is perhaps not too difficult since reliable prior information may be available and, if not, the posterior should be relatively robust to the choice of prior for a low-dimensional parameter. But dealing with the jump component is far less straightforward. Indeed, developing a sound parametric model for $J$, and specifying reasonable priors for the corresponding model parameters, is a non-trivial task: how large and how frequent are the jumps? is the jump size and rate constant in time? etc. Rifo and Torres (2009), for example, in a setting similar to ours in (1), propose a Bayesian model that assumes $J$ is a Poisson process, which is fine for some applications but certainly would not be appropriate for all. Most importantly, the quality of marginal inference on $\theta$ depends on the quality of the posited model for $J$, which is unverifiable. To avoid potential bias from model misspecification, one could go semiparametric, e.g., characterize $J$ by its Lévy measure and put a prior on that, but this severely complicates the posterior computation and, furthermore, the addition of an infinite-dimensional nuisance parameter may affect the efficiency of the marginal inference on $\theta$. Ideally, we would be able to directly attack the interest parameter, to get a posterior distribution for $\theta$ without specifying a model and prior for $J$, computing the full posterior, and putting ourselves at risk of making biased or inefficient inference on $\theta$ as a result of a poor choice of model for $J$.

Towards this, in Section 2, we consider a purposely misspecified model that completely ignores the jumps, basically treating the observations as if they arise from a simple diffusion model. This misspecified model is highly regular and computationally convenient, so if not heavily influenced by misspecification, then perhaps it would suffice for valid inference on $\theta$. A special case of our Theorem 1 says that the misspecified posterior for $\theta$ is asymptotically normal but misspecification causes the center to be off-target and the spread to be too large. Rather than abandon the misspecified model, we propose, in Section 3, to correct for the effects of misspecification, by making two simple location and scale adjustments, both of which rely on an estimator of the quadratic variation of the jump process $J$. We then show, in Theorem 2, that the corresponding modified posterior is asymptotically normal, centered around a consistent estimator of the true volatility, with variance equal to the Cramér–Rao lower bound for optimal/ideal case when there are no jumps, i.e., when the misspecified model is actually correct. As a consequence, no frequentist or proper Bayesian approach – with a parametric or nonparametric model for $J$ – provides asymptotically more efficient inference on $\theta$ than ours. Moreover, our proposed modification is easy to implement and we present some simulation results in Section 4 to illustrate that our posterior credible intervals provide valid uncertainty quantification for the volatility $\theta$.

"Misspecification on purpose" is a general idea which is both practically useful and theoretically interesting, with many potential applications beyond the jump diffusion setup considered here. Our choice to demonstrate the benefits of this general idea in a relatively simple setting is only for the sake of clarity and conciseness. Similar analysis applies in more complex situations but, naturally, the details (work in progress) are more involved and would potentially distract from the general idea.

## 2. A misspecified model

Assume that we observe the continuous-time process $(X_t)$ at $n$ distinct time points $t_1 < \cdots < t_n$, i.e., our observations are $X_{t_1}, \ldots, X_{t_n}$; for notational convenience later on, set $t_0 = 0$ and $X_0 \equiv 0$. For notational simplicity, we will assume