MCMC for Lévy Process Models of Stochastic Volatility

David A. Stephens

Department of Mathematics, Imperial College London
Department of Mathematics and Statistics, McGill University

d.stephens@math.mcgill.ca

Joint work with Matthew Gander and Wing Yip

Outline

- Stochastic Volatility (SV) models in statistical finance
- Lévy processes
- The Barndorff-Nielsen and Shephard (BNS) model
- Inference using MCMC
- Extensions
  - Different marginal models
  - Different correlation structures

Framework

A stochastic process $\{S(t), t > 0\}$ is used to represent the evolution of an asset value through time.

Typically use a process driven by Brownian motion $W(t)$, with drift, satisfying the SDE

$$dS(t) = \mu S(t)dt + \sigma S(t)dW(t)$$

- σ, the volatility, is presumed constant.
- If $X(t) = \log S(t)$, then

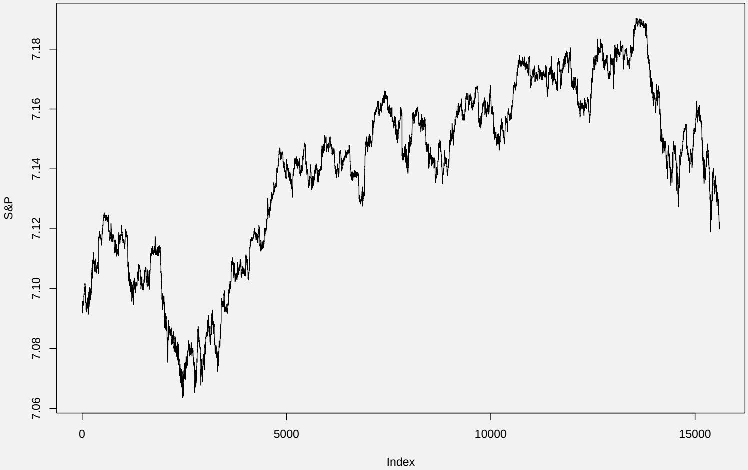
$$dX(t) = \{\mu - \sigma^2/2\}dt + \sigma dW(t).$$

- Typically, the log asset value is observed discretely at intervals of length $\triangle$, say, yielding data $x_0, \ldots, x_T$
- If $Y_t = X_t - X_{t-1}$ then

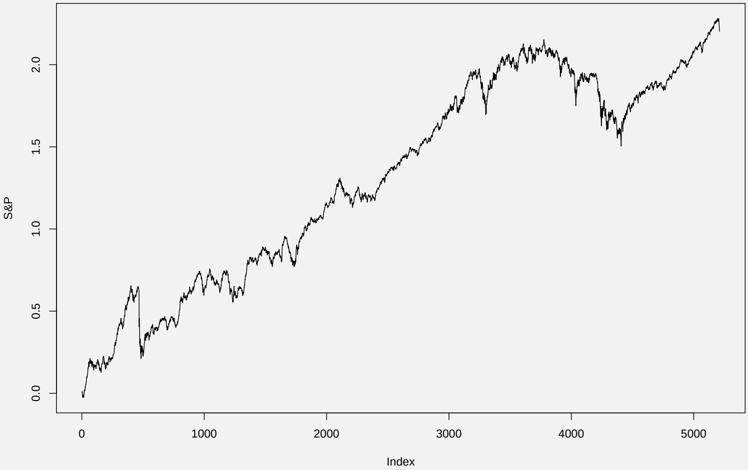
$$Y_t \sim N((\mu - \sigma^2/2)\triangle, \sigma^2\triangle)$$

with $Y_1, \ldots, Y_T$ independent; these are the log returns.

S&P 500 Index (log scale) : Five minute



S&P 500 Index (log scale) : Daily



Deficiencies of constant volatility

The problem is ... this model does not capture observed behaviour.

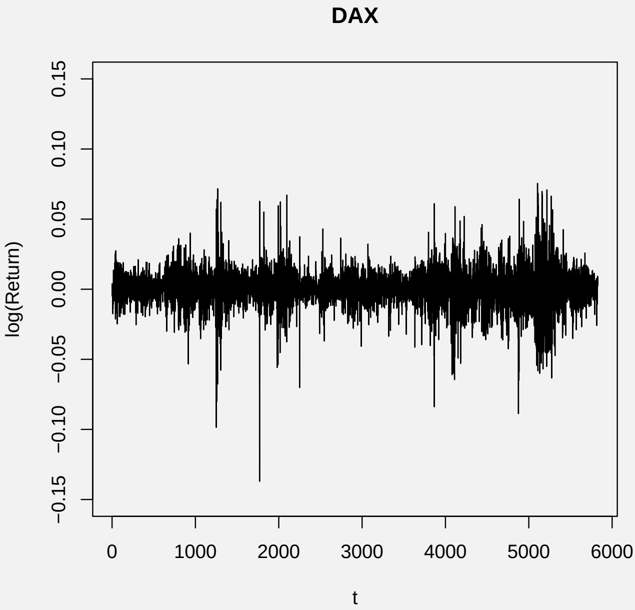
Empirically, using y_t^2 as a proxy for the volatility for time interval $(\triangle(t-1), \triangle t)$, it is apparent that

- volatility is not apparently constant
- volatility exhibits autocorrelation
- marginal distribution of $y_1, \ldots, y_T$ appears leptokurtic (heavy-tailed compared to normal).

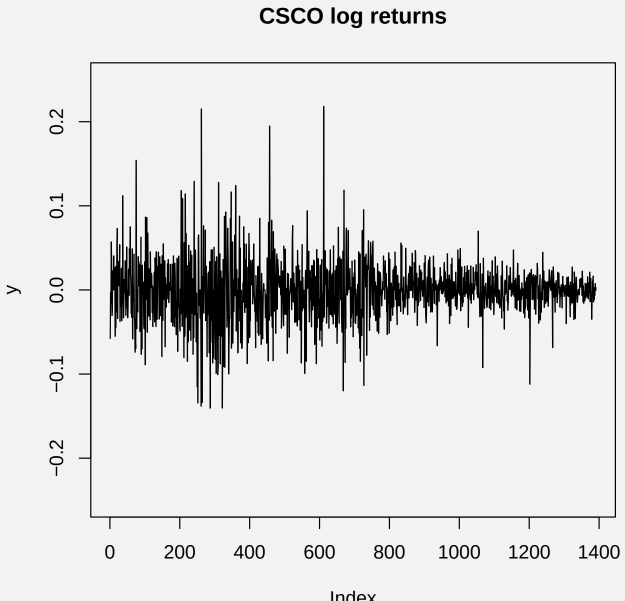
> *... every single assumption underlying the Black-Scholes model is routinely rejected by the type of data that are routinely used in practice.*

Barndorff-Nielsen and Shephard [2001]

DAX Index log returns

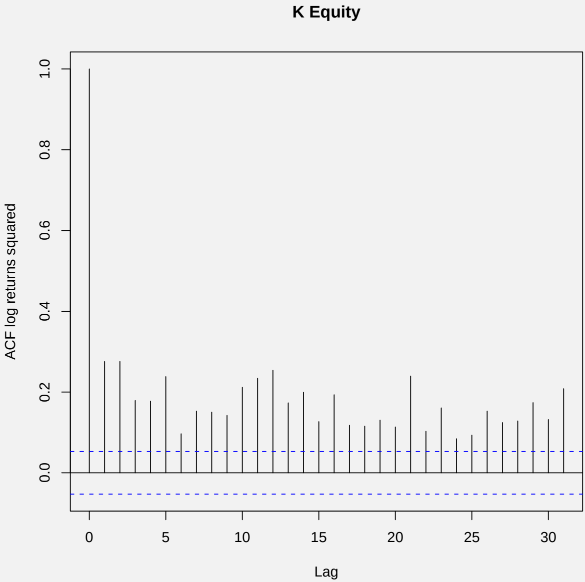


CISCO log returns

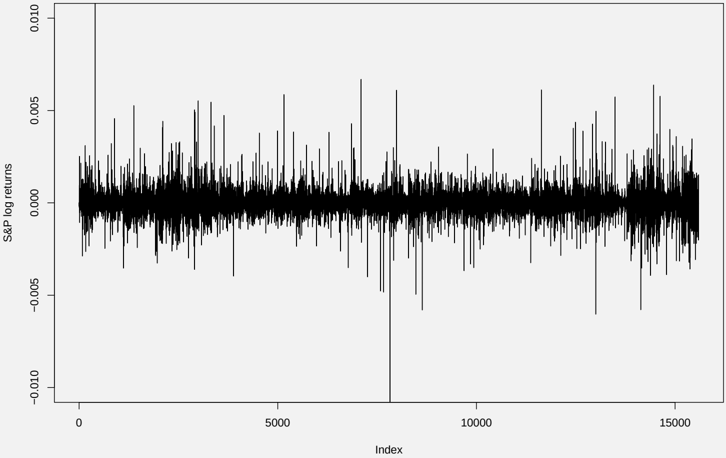


K Equity log returns squared

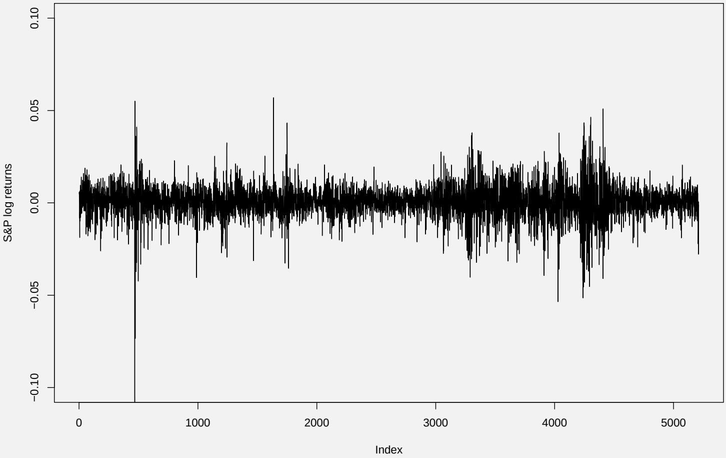
Autocorrelation function (ACF) K Equity stock y_t^2:



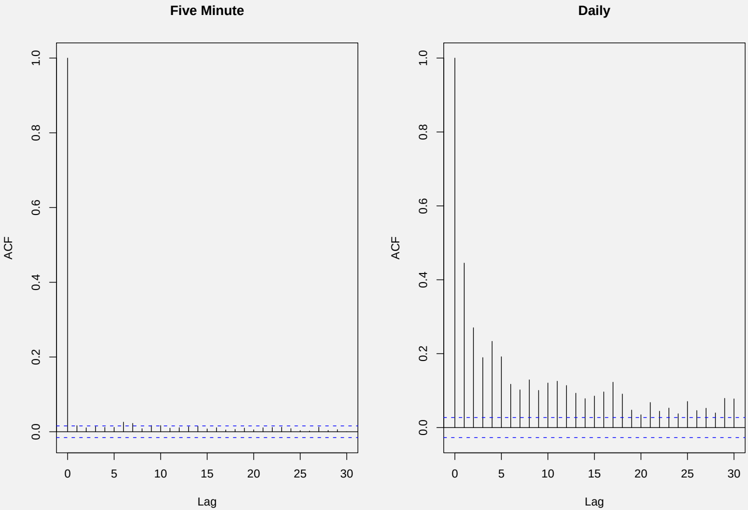
S&P 500 Index (log scale) : Five minute



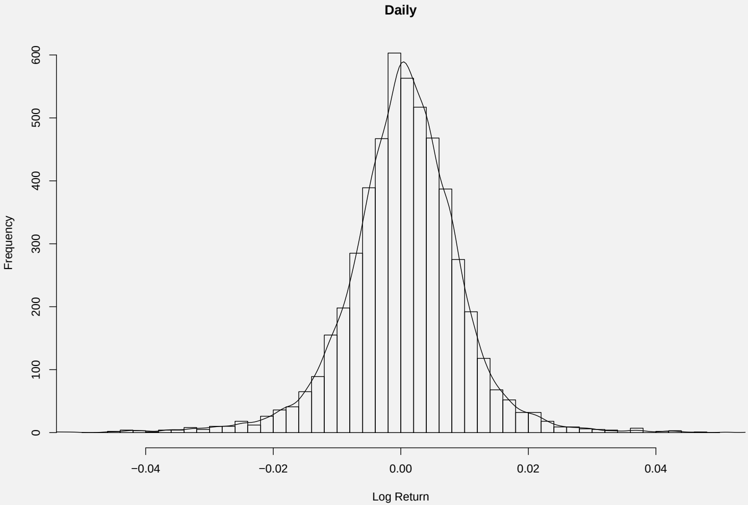
S&P 500 Index (log scale) : Daily



S&P 500 Index : ACF for Squared log-returns



S&P 500 Index : Skewness



Stochastic Volatility: The BNS model

Allow volatility to vary stochastically over time

$$dS(t) = \mu S(t)dt + \sigma(t)S(t)dW(t)$$

- Random volatility increases kurtosis of log returns
- Correlation in volatility process induces correlation in square of log returns

Barndorff-Nielsen and Shephard [2001] (BNS) suggest to make SV process follow the Ornstein-Uhlenbeck (OU) equation

$$d\sigma^2(t) = -\lambda\sigma^2(t)\,dt + dZ(\lambda t)$$

This model induces, for $s, t > 0$,

$$\text{Corr}\left[\sigma^2(s), \sigma^2(s+t)\right] = \exp\{-\lambda t\}.$$

Here, $Z(t), t > 0$ is a Lévy process

Lévy Processes

A Lévy Process, $Z(t), t > 0$ is a continuous time stochastic process such that

1. $Z(0) = 0$.
2. Z has **independent increments**; for $t_0 < t_1 < \ldots < t_n$, the random quantities
$$Z_{t_0}, Z_{t_1} - Z_{t_0}, Z_{t_3} - Z_{t_2}, \ldots, Z_{t_n} - Z_{t_{n-1}}$$
are independent.
3. Z has **stationary increments**; for $t, h > 0$, the distribution of $Z_{t+h} - Z_t$ does not depend on t.
4. Z is **stochastically continuous**; for all $t, h, \varepsilon > 0$,
$$\lim_{h \to 0} P\left[|Z_{t+h} - Z_t| \geq \varepsilon\right] = 0.$$

Lévy Processes and Infinite Divisibility

A distribution, with density $f(x)$ and characteristic exponent $\Psi(s)$, is infinitely divisible iff there exists

- $a \in \mathbb{R}^d$,
- a positive semi-definite quadratic form Q on $\mathbb{R}^d$
- a measure $U(dx)$ on $\mathbb{R}^d/\{0\}$ with density $u(x)$.

such that $\forall s \in \mathbb{R}^d$

$$\Psi(s) = ia^{\mathsf{T}}s + Q(s)/2 + \int_{\mathbb{R}^d} \left[1 - e^{is^{\mathsf{T}}x} + is^{\mathsf{T}}x\mathbf{1}_{\{|x|<1\}}\right] u(x)\,dx$$

and

$$\int_{-\infty}^{\infty} \min\{1, x^2\} u(x)\,dx < \infty.$$

Lévy Processes and Infinite Divisibility

The identity above is called the Lévy-Khintchine formula

- $U(dx)$ is the Lévy measure of f
- $u(x)$ is the Lévy density of f
- $Q(s)$ is the Gaussian coefficient.
- a 1-1 correspondence between ID distributions and Lévy processes
- $f(x)$ is the marginal law of Z.

In

$$d\sigma^2(t) = -\lambda\sigma^2(t)dt + dZ(\lambda t)$$

we interpret the $dZ(t)$ term as the (random) change in Z at instant t.

If the $Q(s) \equiv 0$, then $Z(t)$ only changes in jumps; need positive jumps.

BNS OU-SV model

In the BNS OU-SV model:

- Z is termed the Background Driving Lévy process (BDLP)
- Z is a pure jumps process, and is non-decreasing
- Z has an associated marginal law (given by the Lévy-Khintchine theorem) that **does not depend on λ**.
- Parameter λ controls rate of jumps of $\sigma^2(t)$.

Strategy is to pick appropriate Z to induce appropriate marginal law for $\sigma^2(t)$.

No theoretical reason to prefer one model over another - should make assessment statistically.

Lévy Measures for $\sigma^2(t)$ and $Z(t)$

Denote by

- $u(x)$ the Lévy density for the (ID) law of $\sigma^2(t)$
- $w(x)$ the Lévy density for the marginal law of $Z(1)$.
- Under the BNS OU-SV specification

$$w(x) = -u(x) - x\frac{du(x)}{dx}$$

- Tail Mass function (TMF)

$$W^+(x) = \int_x^\infty w(y)\,dy = xu(x)$$

- Inverse Tail mass function (ITMF)

$$W^{-1}(x) = \inf\{y > 0 : W^+(y) \leq x\}$$

OU Equation: Solution and Simulation

Solution to OU equation can be written via a stochastic integral wrt BDLP

$$\sigma^2(t) = e^{-\lambda t}\sigma^2(0) + e^{-\lambda t}\int_0^t e^{\lambda s} dZ(\lambda s)$$

The Ferguson and Klass [1972] infinite series representation of $Z(t)$ yields a means of

- simulating $Z(t)$
- performing inference about $\sigma^2(t)$

We use the Inverse-Lévy Method: other simulation methods exist.

OU Equation: Solution and Simulation

We have

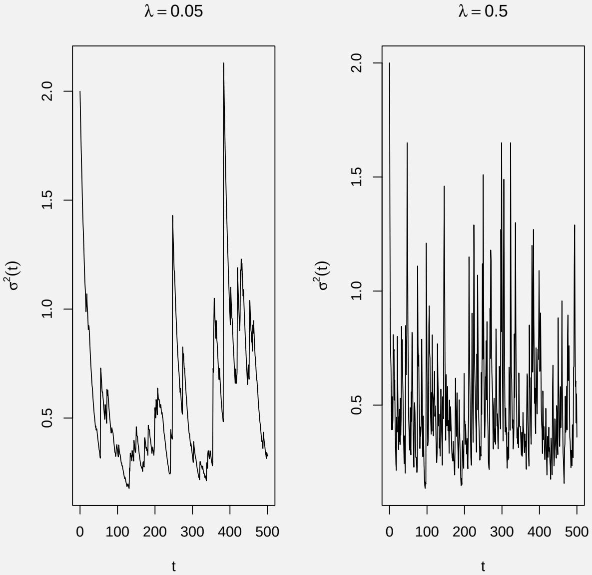
$$\int_0^{\Delta} f(s)\, dZ(s) \stackrel{\mathcal{L}}{=} \sum_{j=1}^{\infty} f(\Delta r_j)\, W^{-1}\left(\frac{a_j}{\Delta}\right)$$

where

- $a_1 < a_2 < \ldots$ are a sequence of event times in a standard Poisson process
- r_j are independent Uniforms
- W^{-1} is the ITMF defined earlier.

Auxiliaries (a_j, r_j), $j = 1, 2, \ldots,$ facilitate implementation.

Examples: Realizations of $\{\sigma^2(t)\}$



BNS OU-SV: Bayesian Inference

With this model for $\{\sigma^2(t)\}$, we can carry out Bayesian inference in light of the observed data

For the likelihood, similarly to the constant volatility case

$$Y_t \sim N((\mu - \sigma_t^2/2)\triangle, \sigma_t^2\triangle)$$

with $Y_1, ..., Y_T$ independent, where

$$\sigma_t^2 = \int_0^{\triangle t} d\sigma^2(s) - \int_0^{\triangle(t-1)} d\sigma^2(s)$$

is the discretely observed volatility.

BNS OU-SV: Bayesian Inference

In the BNS OU-SV models, we have the following parameters to make inference about

- λ, autocorrelation decay parameter.
- θ, the parameters of the marginal law of $\sigma^2(t)$.
- the list of latent variables that form the (Ferguson & Klass) representation of the BDLP, that is

$$(a_1, a_2, \ldots) \qquad (r_1, r_2, \ldots).$$

- $\sigma^2(0)$, initial value of volatility process.
- More sophisticated model components can be added eg leverage terms.

BNS OU-SV: MCMC

Several MCMC algorithms and series representations exist and are in use: see

- Roberts et al. [2004]
- Griffin and Steel [2005]
- ...

MCMC issues:

- parameterization: non-centering
- dependence: overconditioning

BNS OU-SV: MCMC

Useful speedups :

- Joint updates for auxiliaries and parameters
- Targeted Proposals : target updates of auxiliaries (a_j, r_j) in areas of poor fit to data under current state. Interval i
  - has return y_i
  - has collection of auxiliaries $\theta^{(i)} = (a^{(i)}, r^{(i)})$
  - has a discretely observed volatility $\sigma_i^2 \equiv \sigma_i^2(\theta^{(i)})$ given any set of parameter values
  - $\therefore$ target update of $\theta^{(i)}$ where residual $(y_i - \mu_i)/\sigma_i$ is largest, that is, select i with probability proportional to

$$\exp\{(y_i - \mu_i)^2/\sigma_i^2\}$$

and correct in Metropolis-Hastings ratio.

BNS OU-SV: Marginal Laws

Most interest has focussed on the Generalized Inverse Gaussian (GIG) family of distributions as suitable as marginal laws for $\sigma^2(t)$.

Specifically, MCMC inference has been implemented for a Gamma marginal law; the ITMF takes a particularly simple form, and the infinite series in the series representation of Z reduces to a finite one.

This simplifies the computation significantly, although the MCMC is still not trivial.

Deficiencies of Gamma BNS OU-SV model

The problem is ... this model may not capture observed behaviour.

- no particular theoretical support for a Gamma marginal
- autocorrelation decay is not always exponential
- volatility process is perhaps non-stationary

Need extension to full GIG marginal model: for parameters $\gamma \in \mathbb{R}$ and $\nu, \alpha > 0$, the $GIG(\gamma, \nu, \alpha)$ pdf takes the form

$$f(x) = \frac{(\alpha/\nu)^{\gamma}}{2K_{\gamma}(\nu\alpha)} x^{\gamma-1} \exp\left\{-\frac{1}{2}\left(\nu^2 x^{-1} + \alpha^2 x\right)\right\}$$

for $x > 0$, where K_{ν} is a modified Bessel function of the third kind.

TMF/ITMF

For $GIG(\gamma, \nu, \alpha)$

- Lévy density

$$u(x) = \frac{1}{x}\exp\left(-\frac{\alpha^2 x}{2}\right) \left\{\frac{1}{2}\int_0^\infty \exp\left(-\frac{x\xi}{2\nu^2}\right) g_\gamma(\xi)\, d\xi + \max(0,\gamma)\right\}$$

where

$$g_\gamma(x) = \frac{2}{x\pi^2}\left\{J^2_{|\gamma|}\left(\sqrt{x}\right) + N^2_{|\gamma|}\left(\sqrt{x}\right)\right\}^{-1}$$

and $J_{|\nu|}$ and $N_{|\nu|}$ are Bessel functions of the first and second kind

- TMF $W^{+}(x) = xu(x)$
- ITMF

$$W^{-1}_{\gamma,\nu,\alpha}(x) = z,$$

where z satisfies

$$x = \left\{ \frac{1}{2} \int_0^\infty \exp\left(-\frac{z\xi}{2\nu^2}\right) g_\gamma(\xi)\, d\xi + \max(0,\gamma) \right\} \times \exp\left(-\frac{\alpha^2 z}{2}\right)$$

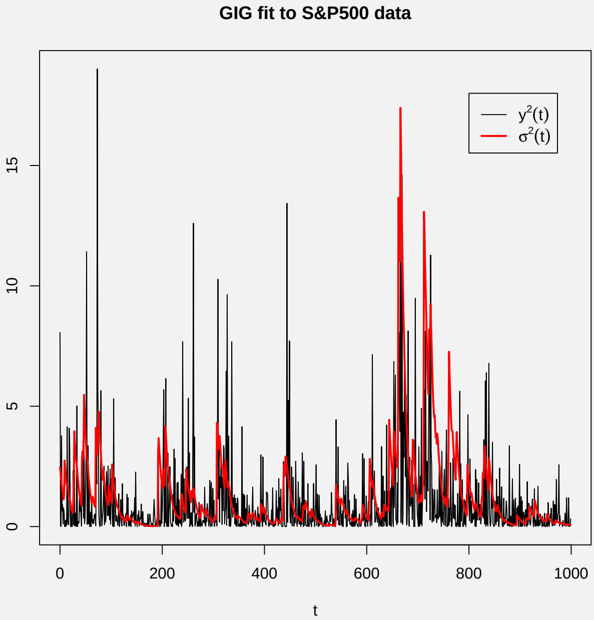
Evaluating the TMF/ITMF to arbitrary accuracy is possible using numerical integration.

GIG Special cases

Form of $GIG(\gamma, \nu, \alpha)$ distribution	Distribution
$GIG\left(\nu, 0, \sqrt{2\alpha}\right)$	$Ga(\nu, \alpha)$
$GIG(1, \nu, \alpha)$	$PH(\nu, \alpha)$
$GIG\left(-\nu, \sqrt{2\alpha}, 0\right)$	$IGa(\nu, \alpha)$
$GIG\left(-\frac{1}{2}, \nu, \alpha\right)$	$IG(\nu, \alpha)$

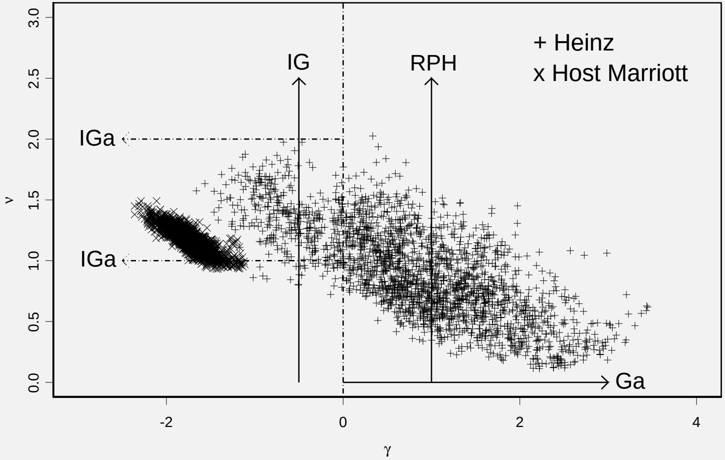
Ga is Gamma, *PH* is positive hyperbolic, *IGa* is Inverse Gamma, *IG* is Inverse Gaussian

GIG Fit to S & P 500 Data



Other posterior analysis available

Posterior samples in the (γ, ν) plane for the Heinz and Host Marriott data sets.



BNS OU-SV-GIG: Summary

- Desirable to be able to compare performance for across marginal models.
- Forecasting/pricing routinely possible; gives framework for model comparison.
- Analysis suggests that Inverse Gamma marginal model preferable for option pricing for these share series.
- Other model assessment methods (eg posterior predictive) for as yet unobserved returns support the three-parameter GIG model for these data; Gamma does OK.

Extension: More General Correlation Structure

The exponential decay in the autocorrelation provided by the BNS OU-SV model is limited.

- One nice method for generalizing autocorrelation is to use superposition
  - Set
$$\sigma^2(t) = \sum_{k=1}^{K} w_k \sigma_k^2(t)$$
    where $\sigma_k^2(t)$ is BNS OU with parameter λ_k.
  - $K = 2$ readily implemented
- Can also use the approaches of Wolpert and Taqqu [2005] to construct volatility processes with richer autocorrelation.
  - Generally more computational expense, but still feasible to perform MCMC.

The Wolpert and Taqqu Approach

The BNS OU-SV process can be written

$$\begin{aligned}
\sigma^2(t) &= \int_{-\infty}^{t} f_2(\lambda,t,s)\,dZ(\lambda s) \\
&= \int_0^{\infty} f_1(\lambda,t,s)\,dZ(\lambda s) + \int_0^{t} f_2(\lambda,t,s)\,dZ(\lambda s) \\
&= e^{-\lambda t}\sigma^2(0) + e^{-\lambda t}\int_0^t e^{\lambda s}\,dZ(\lambda s),
\end{aligned}$$

where the two terms in the second equation are independent, and

$$f_1(\lambda,t,s) = e^{-\lambda(t+s)} \qquad f_2(\lambda,t,s) = e^{-\lambda(t-s)}.$$

Changing the correlation structure

The idea is to change the choice of f_1 and f_2 that appear in the form

$$\sigma^2(t) = \int_0^\infty f_1(\lambda, t, s)\, dZ(\lambda s) + \int_0^t f_2(\lambda, t, s)\, dZ(\lambda s)$$

so that the correlation structure for $\sigma^2(t)$ changes, but the process remains stationary.

Wolpert and Taqqu [2005] give details of how to do this.

For these models, computing the required stochastic integrals is harder, but still possible using series representations.

Alternative Models

Wolpert and Taqqu [2005] adopt the moving average model

$$\sigma^2(t) = \int_{-\infty}^{t} h_1(t-s)dZ(s)$$

where $h_1(t-s) \geq 0$ if $t > s$.

Because of the timing change on the Lévy process, we may not be able to make the marginal law of $\sigma^2(t)$ vary independently of the autocorrelation.

The correlation at lag t for this model is

$$\rho(t) = \text{Corr}\left[\sigma^2(t_0), \sigma^2(t_0+t)\right] = \frac{\int_0^\infty h_1(|t|+s)h_1(s)ds}{\int_0^\infty h_1^2(s)ds}.$$

Power Decay and fOUL model

- Power Decay model

$$h_1(x) = \frac{1}{(\alpha + \beta |x|)^{\lambda}}$$

which, for large lags t, yields an ACF that decays like $t^{-\lambda}$. If $\lambda \to 1$, then model exhibits (quasi) long range dependence.

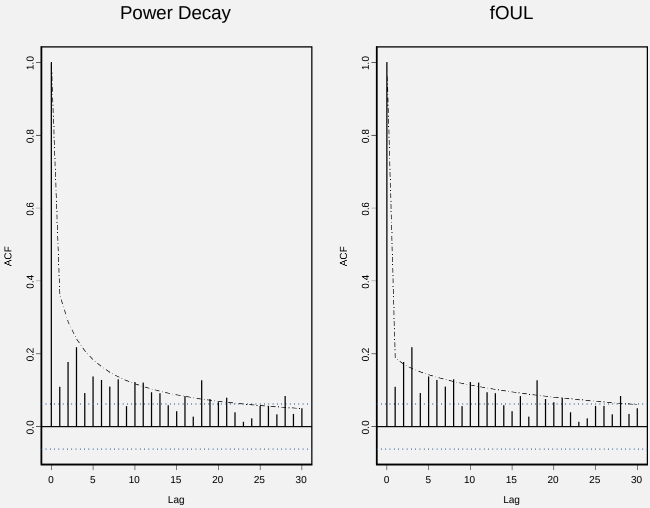
- Fractional OU Lévy Model

$$h_1(x) = \sqrt{2\lambda}\frac{\lambda^{\kappa-1}}{\Gamma(\kappa)}x^{\kappa-1}e^{-\lambda x} \quad x \geq 0, \text{ zero otherwise}$$

which has finite variance (and short memory) if $\kappa > 1/2$.

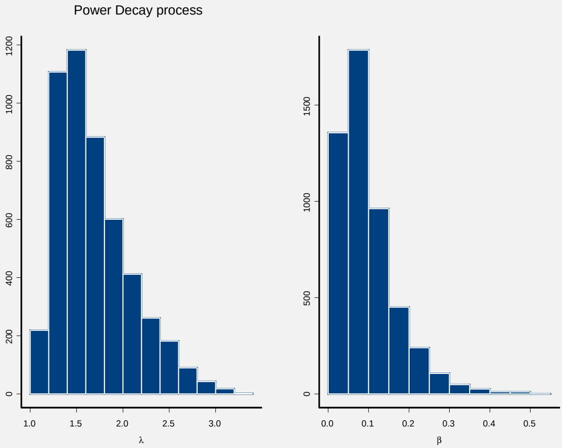
Example: Power decay and fOUL model

ACF of the square of the log returns of S&P 500 data and theoretical ACF of fitted Power Decay and fOUL processes.



S&P 500 data: No long memory ?

Posterior for λ has most mass away from $\lambda = 1$.



Extension: Changing the Observation Equation

Work in Progress: collaboration with Wing Yip

Motivation: To use a more general observation model to better capture observed market behaviour.

This can be achieved by using a Lévy process with jump components rather than a Brownian motion to drive the observation equation.

Parallel motivation: Wing Yip's employers (RBS) don't care too much about Bayesian inference, MCMC implementation, representations of Lévy processes ...

The Variance Gamma Model

The Variance Gamma model for returns data is based on a time-changed drifting Brownian motion

$$X(t) = \mu G(t) + \sigma W(G(t))$$

where $\{G(t)\}$ is a (one-parameter) Gamma process

- a pure-jumps non-decreasing Lévy process
- independent and stationary increments are Gamma distributed
- an infinite activity Lévy process.

Under this model, log returns are independent, identically distributed variates with tractable density that exhibits skewness and kurtosis.

The Variance Gamma Model: Option Pricing

Under the VG model, can compute option prices in closed form.

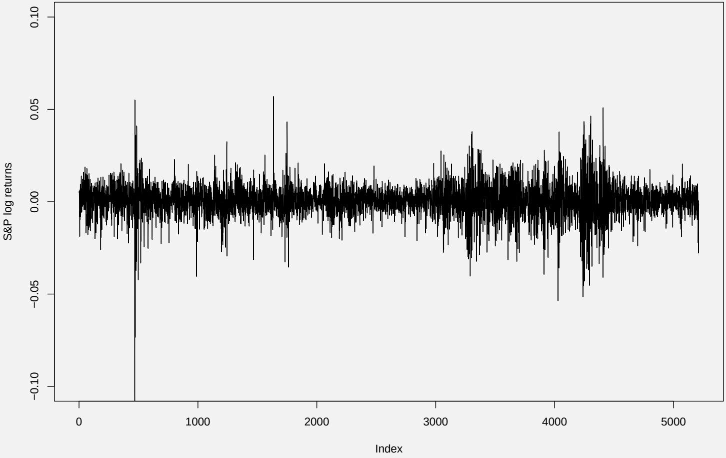
The RBS motivation is that parameter estimates derived from returns data do not necessarily correspond to estimates derived from OLS procedure on logged option prices.

Thus it may be possible to exploit mis-pricing.

However, as specified, this model is not appropriate ...

Recall: S&P 500 Index (log scale)

Distinct volatility clustering:



The BNS VG-OU-SV-GIG Model

Need to combine the VG and the SV model:

$$X(t) = \mu G(t) + \sigma(t) W(G(t))$$

or possibly

$$X(t) = \mu G(t) + \sigma(G(t)) W(G(t))$$

where $\sigma^2(t)$ follows the BNS OU-SV specification.

- Inference for this model is in principle not radically more difficult than for the BNS OU-SV model
- Other time changes can be used
- Analytic option pricing results often available for such models.
- Classical estimation achieved using transform methods
- Discrete-time approximations in simpler models already implemented in MCMC

The General Model Form

Li et al. [2004]: Model utilized:

$$\begin{pmatrix} dY(t) \\ d\sigma^2(t) \end{pmatrix} = \begin{pmatrix} \mu \\ \lambda(\theta - \sigma^2(t)) \end{pmatrix} dt + \sigma(t) \begin{pmatrix} 1 & 0 \\ \rho\tau_\sigma & \sqrt{(1-\rho^2)}\tau_\sigma \end{pmatrix} \begin{pmatrix} dW_y(t) \\ dW_\sigma(t) \end{pmatrix} + \begin{pmatrix} dJ_y(t) \\ dJ_\sigma(t) \end{pmatrix}$$

Particular Case: SV-VG

Discrete Approximation:

$$Y_{t+1} = Y_t + \mu\triangle + \sigma_t\sqrt{\triangle}\varepsilon^y_{t+1} + J^y_{t+1}$$

$$\sigma^2_{t+1} = \sigma^2_t + \lambda(\theta - \sigma^2_t)\triangle + \tau_\sigma\sigma_t\sqrt{\triangle}\varepsilon^\sigma_{t+1}$$

with

$$\varepsilon^y_{t+1}, \varepsilon^\sigma_{t+1} \sim N(0,1) \qquad Corr[\varepsilon^y_{t+1}, \varepsilon^\sigma_{t+1}] = \rho$$

and

$$J^y_{t+1} = \alpha G_{t+1} + \sigma\sqrt{G_{t+1}}\varepsilon^J_{t+1}$$

with

$$G_{t+1} \sim Gamma\left(\frac{\triangle}{\nu}, \nu\right) \qquad \varepsilon^J_{t+1} \sim N(0,1)$$

MCMC possible - many auxiliaries ...

Conclusions

- Can extend the BNS-SV models in different directions
- Going beyond the Gamma marginal is useful
- Going beyond exponential decay in correlation is vital
- Jump components to driving processes probably needed
- Sequential methods probably offer more attractive realtime analysis

References

O E Barndorff-Nielsen and N Shephard. Non-Gaussian Ornstein-Uhlenbeck based models and some of their uses in financial economics. *Journal of the Royal Statistical Society, Series B*, 63:167–241, 2001.

T S Ferguson and M J Klass. A representation of independent increment processes without Gaussian components. *Annals of Mathematical Statistics*, 43:1634–1643, 1972.

J E Griffin and M F J Steel. Inference with non-Gaussian Ornstein-Uhlenbeck processes for stochastic volatility. *Journal of Econometrics*, 2005. In press.

H. Li, M. Wells, and L Yu. A MCMC analysis of time-changed Lévy processes of stock return dynamics. Technical report, Cornell University, 2004. Technical Report.

G Roberts, O Papaspiliopoulos, and P Dellaportas. Bayesian inference for non-Gaussian Ornstein-Uhlenbeck stochastic volatility processes. *Journal of the Royal Statistical Society, Series B*, 66:369–393, 2004.

R L Wolpert and M S Taqqu. Fractional Ornstein-Uhlenbeck Lévy processes and the telecom process: Upstairs and downstairs. *Signal Processing*, 85(8): 1523–1545, 2005.