Sample path asymmetries in random processes driven by a second-order Lévy motion

Anastassia Baxevani
Department of Mathematics and Statistics
University of Cyprus
with Podgórski, K and Wegener, J.

August 28, 2014

Outline

A-B-C of good modeling

A-B-C of good modeling

- A) Data usually exhibit certain features like asymmetries ...

A-B-C of good modeling

- A) Data usually exhibit certain features like asymmetries ...
- B) so we need models that can reproduce these features...

A-B-C of good modeling

- A) Data usually exhibit certain features like asymmetries ...
- B) so we need models that can reproduce these features...
- C) and then we need some characteristics/statistics that can describe these features.

Gaussian models not always sufficient

Some good reasons for using Gaussian models:

Gaussian models not always sufficient

Some good reasons for using Gaussian models:

- Fully characterized by the first two moments

Gaussian models not always sufficient

Some good reasons for using Gaussian models:

- Fully characterized by the first two moments
- Likelihood accessible

Gaussian models not always sufficient

Some good reasons for using Gaussian models:

- Fully characterized by the first two moments
- Likelihood accessible
- Conditional expectation is linear

Gaussian models not always sufficient

Some good reasons for using Gaussian models:

- Fully characterized by the first two moments
- Likelihood accessible
- Conditional expectation is linear
- Stability under linear combinations, marginalization and conditioning

Gaussian models not always sufficient

Some good reasons for using Gaussian models:

- Fully characterized by the first two moments
- Likelihood accessible
- Conditional expectation is linear
- Stability under linear combinations, marginalization and conditioning

Some good reasons for NOT using Gaussian models:

Gaussian models not always sufficient

Some good reasons for using Gaussian models:

- Fully characterized by the first two moments
- Likelihood accessible
- Conditional expectation is linear
- Stability under linear combinations, marginalization and conditioning

Some good reasons for NOT using Gaussian models:

- Always symmetric since they depend solely on their covariance function

Gaussian models not always sufficient

Some good reasons for using Gaussian models:

- Fully characterized by the first two moments
- Likelihood accessible
- Conditional expectation is linear
- Stability under linear combinations, marginalization and conditioning

Some good reasons for NOT using Gaussian models:

- Always symmetric since they depend solely on their covariance function
- Have difficulty accounting for unusually extreme values in the data

Gaussian models not always sufficient

Some good reasons for using Gaussian models:

- Fully characterized by the first two moments
- Likelihood accessible
- Conditional expectation is linear
- Stability under linear combinations, marginalization and conditioning

Some good reasons for NOT using Gaussian models:

- Always symmetric since they depend solely on their covariance function
- Have difficulty accounting for unusually extreme values in the data

So we need to go beyond the Gaussian world!

Leaving the Gaussian world, but not too far...

Playing with Gaussian RMs

- Transforming
- Thresholding: Excursion sets
- Truncating: Truncated Gaussian or transformed Gaussian RMs
- Conditioning: Skew-normal RMs

does not produce all types of asymmetries met in the real world.

Leaving the Gaussian world, but not too far...

Playing with Gaussian RMs

- Transforming
- Thresholding: Excursion sets
- Truncating: Truncated Gaussian or transformed Gaussian RMs
- Conditioning: Skew-normal RMs

does not produce all types of asymmetries met in the real world.

- Laplace moving average models are formulated as convolutions of Laplace noise and some deterministic kernel

Leaving the Gaussian world, but not too far...

Playing with Gaussian RMs

- Transforming
- Thresholding: Excursion sets
- Truncating: Truncated Gaussian or transformed Gaussian RMs
- Conditioning: Skew-normal RMs

does not produce all types of asymmetries met in the real world.

- Laplace moving average models are formulated as convolutions of Laplace noise and some deterministic kernel
  - Flexible and rich: 4 parameters + kernel

Leaving the Gaussian world, but not too far...

Playing with Gaussian RMs

- Transforming
- Thresholding: Excursion sets
- Truncating: Truncated Gaussian or transformed Gaussian RMs
- Conditioning: Skew-normal RMs

does not produce all types of asymmetries met in the real world.

- Laplace moving average models are formulated as convolutions of Laplace noise and some deterministic kernel
  - Flexible and rich: 4 parameters + kernel
  - Can account for both geometrical asymmetries and occasional highly extreme events

Outline

Moving averages through stochastic integration

- Standard construction of stochastic integrals with deterministic kernels :

$$X(\tau) = \int_{\mathbb{R}} f(\tau - x) d\Lambda(x).$$

Moving averages through stochastic integration

- Standard construction of stochastic integrals with deterministic kernels :

$$X(\tau) = \int_{\mathbb{R}} f(\tau - x) d\Lambda(x).$$

- $\Lambda(x)$- Laplace motion has 4 parameters for location (δ usually set to zero), symmetry (μ), scale (σ) and shape (ν)

Moving averages through stochastic integration

- Standard construction of stochastic integrals with deterministic kernels :

$$X(\tau) = \int_{\mathbb{R}} f(\tau - x) d\Lambda(x).$$

- $\Lambda(x)$- Laplace motion has 4 parameters for location (δ usually set to zero), symmetry (μ), scale (σ) and shape (ν)
- Covariance :

$$r(\tau) = \frac{\sigma^2 + \mu^2}{\nu} \int f(x - \tau) f(x) dx$$

Moving averages through stochastic integration

- Standard construction of stochastic integrals with deterministic kernels :

$$X(\tau) = \int_{\mathbb{R}} f(\tau - x) d\Lambda(x).$$

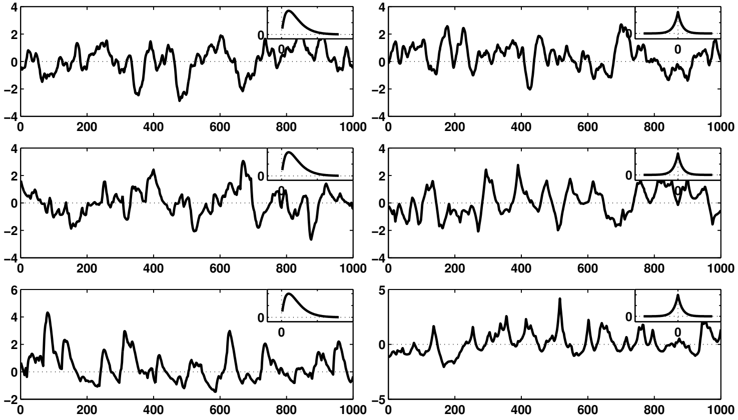
- $\Lambda(x)$- Laplace motion has 4 parameters for location (δ usually set to zero), symmetry (μ), scale (σ) and shape (ν)
- Covariance :

$$r(\tau) = \frac{\sigma^2 + \mu^2}{\nu} \int f(x - \tau) f(x) dx$$

- Spectral density:

$$R(\omega) = \frac{\sigma^2 + \mu^2}{\nu} \mathcal{F}f(\omega) \mathcal{F}f(-\omega)$$

Laplace moving average trajectories



Outline

Fitting the model

- The problem of fitting the model is two folded:

Fitting the model

- The problem of fitting the model is two folded:
  - Fit the kernel

Fitting the model

- The problem of fitting the model is two folded:
  - Fit the kernel
  - Fit the Laplace noice

Symmetric kernel case

Symmetric kernel case

- A non-parametric approach is to estimate $\widehat{f}$ by

$$\widehat{f}(x) = \mathcal{F}^{-1}\sqrt{\widehat{S}(\omega)},$$

where $\widehat{S}(\omega)$ is an estimate of spectrum

Non-symmetric kernels – measures of asymmetries

- There is no general non-parametric approach to the problem.

Non-symmetric kernels – measures of asymmetries

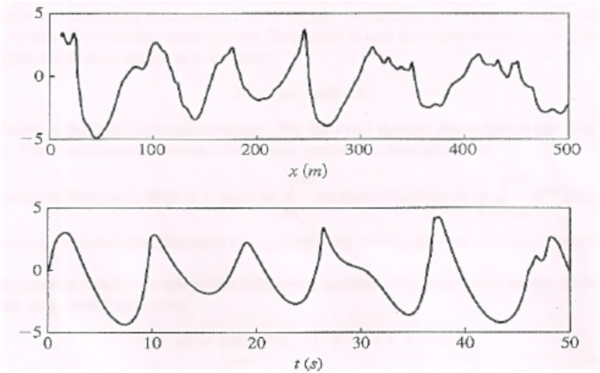
- There is no general non-parametric approach to the problem.
  - Various parametrized families of kernels are available

Non-symmetric kernels – measures of asymmetries

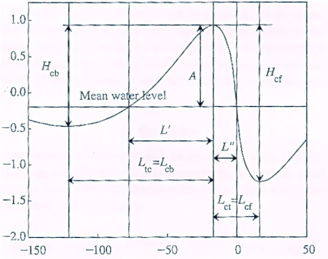
- There is no general non-parametric approach to the problem.
    - Various parametrized families of kernels are available
- But what measures/ methods can we use?

Measures of asymmetry

A non-linear model for sea surface in space and time



Measures of asymmetry in observed records



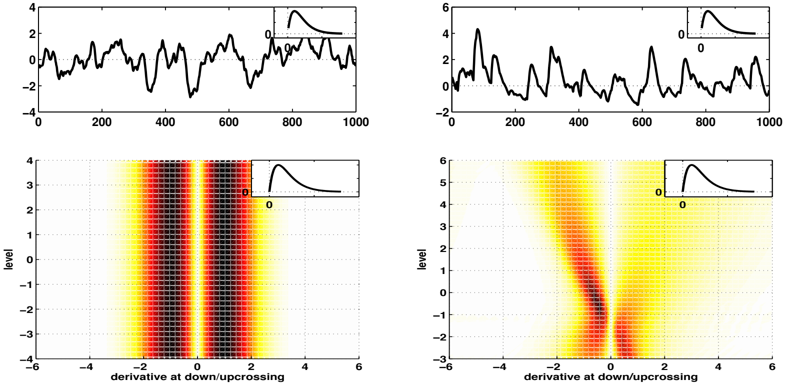
- Measures of asymmetries through crossing distributions of the characteristics:

$$\lambda_{MK} = L'/L", \quad \lambda_{NLS} = L_{cb}/L_{cf}$$

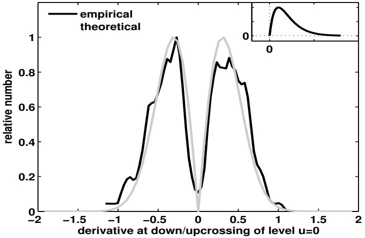
or if L_x is a derivative at the mean water level:

$$\lambda_{AL} = -L_x(x_{down})/L_x(x_{up})$$

Tilting of trajectories



Tilting of trajectories



So measures of asymmetry should involve moments of joint distribution of the process and its derivative

Measuring tilting using Rice's formula

- $N(T, A)$ – “number” of times the field X takes value u in $[0, T]$ and at the same time has a property A
- For ergodic stationary processes

$$\lim_{T\to\infty} \frac{N(T,A)}{N(T)} = \frac{\mathbb{E}\left[\{X \in A\}|\dot{X}(0)|\middle|X(0) = u\right]}{\mathbb{E}\left[|\dot{X}(0)|\middle|X(0) = u\right]},$$

- The right hand side represents the biased sampling distribution when sampling is made over the u-level contour $\mathcal{C}_u = \{\boldsymbol{\tau} : X(\boldsymbol{\tau}) = u\}$

Measures of tilting

Measures of tilting should involve derivatives of the LMA process. We propose:

- $$\rho_3 = \frac{\mathbb{E}\dot{X}^3}{\mathbb{E}^{3/2}\dot{X}^2},$$

  - skewness of the derivative process. If process has symmetric distribution, this measure fails to capture any asymmetries in the records!

Measures of tilting

Measures of tilting should involve derivatives of the LMA process. We propose:

- $$\rho_3 = \frac{\mathbb{E}\dot{X}^3}{\mathbb{E}^{3/2}\dot{X}^2},$$

  - skewness of the derivative process. If process has symmetric distribution, this measure fails to capture any asymmetries in the records!

- $$\rho_{3,1} = \frac{\mathbb{E}(\dot{X}^3 X)}{\mathbb{E}^{3/2}\dot{X}^2 \mathbb{E}^{1/2} X^2} \text{ or } r_{3,1} = \frac{\mathbb{E}(\dot{X}^3 X)}{\mathbb{E}^2 X^2}$$

The Gamma kernel

- $$f(x;\tau,\beta) = \frac{2^{\tau-1/2}}{\sqrt{\Gamma(2\tau-1)}\beta^{\tau-1/2}} x^{\tau-1} e^{-x/\beta} \qquad x \geq 0,\ \tau > 0,\ \beta > 0,$$

  - for $\tau = 1$ exponential for $\tau \to \infty$ Gaussian case.

The Gamma kernel

- $$f(x;\tau,\beta) = \frac{2^{\tau-1/2}}{\sqrt{\Gamma(2\tau-1)}\beta^{\tau-1/2}} x^{\tau-1} e^{-x/\beta} \qquad x \geq 0,\ \tau > 0,\ \beta > 0,$$
  - for $\tau = 1$ exponential for $\tau \to \infty$ Gaussian case.
- $$r_{31} = 3\nu \frac{4^{2(2-\tau)}(\tau-1)\Gamma(4\tau-6)}{\beta^4 \Gamma^2(2\tau-1)}, \qquad \tau > 3/2 .$$

The Gamma kernel

- $$f(x;\tau,\beta) = \frac{2^{\tau-1/2}}{\sqrt{\Gamma(2\tau-1)}\beta^{\tau-1/2}} x^{\tau-1} e^{-x/\beta} \qquad x \geq 0,\ \tau > 0,\ \beta > 0,$$

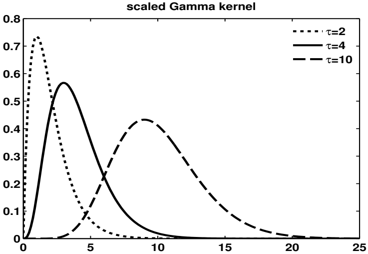
  \- for $\tau = 1$ exponential for $\tau \to \infty$ Gaussian case.

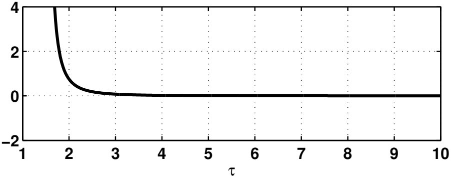
- $$r_{31} = 3\nu \frac{4^{2(2-\tau)}(\tau-1)\Gamma(4\tau-6)}{\beta^4\Gamma^2(2\tau-1)}, \qquad \tau > 3/2\,.$$

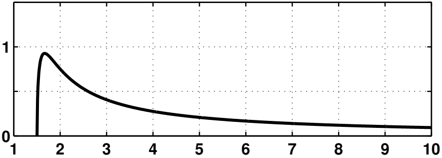
τ	$E(\hat{\tau})$	$Std(\hat{\tau})$
2	2.0245	.0549
4	4.1367	.4636
10	11.628	4.364

Monte-Carlo estimation of τ, for $\nu = 1$ and $\mu = 0$. Hundred Monte-Carlo replicates of sample size of 300000, with step 0.1.

The Gamma kernel , cont.

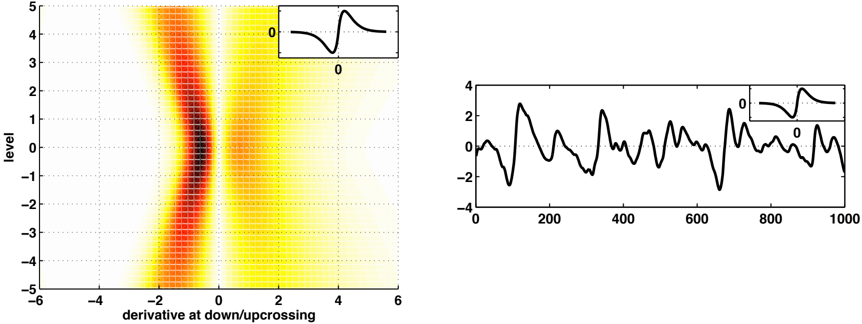






$\beta = 1$ and tilting measures r_{31} (*top*), and ρ_{31} (*bottom*) as functions of τ.

Tilting o LMA with odd kernels, (always symmetric marginals)



Marginal distribution of LMA always symmetric, but trajectories may exhibit some asymmetric features due to asymmetry of kernel which may be captured by the third moments of the derivative.

Distributional parameters

Distributional parameters

- Method of Moments

Distributional parameters

- Method of Moments
- Maximum Likelihood Estimation

Distributional parameters

- Method of Moments
- Maximum Likelihood Estimation
- Maximum Likelihood Estimation through EM algorithm

Highlights

- Non-Gaussian stochastic fields are proposed that can be suitable for modeling environmental data.

Highlights

- Non-Gaussian stochastic fields are proposed that can be suitable for modeling environmental data.
- The models are introduced by means of integrals with respect to independently scattered stochastic measures that have generalized Laplace distributions.

Highlights

- Non-Gaussian stochastic fields are proposed that can be suitable for modeling environmental data.
- The models are introduced by means of integrals with respect to independently scattered stochastic measures that have generalized Laplace distributions.
- Resulting stationary second order processes have, as opposed to their Gaussian counterpart, a possibility of accounting for asymmetry and heavier tails.

Highlights

- Non-Gaussian stochastic fields are proposed that can be suitable for modeling environmental data.
- The models are introduced by means of integrals with respect to independently scattered stochastic measures that have generalized Laplace distributions.
- Resulting stationary second order processes have, as opposed to their Gaussian counterpart, a possibility of accounting for asymmetry and heavier tails.
- Despite this greater flexibility the discussed models still share a lot of spectral properties with Gaussian processes having the latter as a special case.

Highlights

- Non-Gaussian stochastic fields are proposed that can be suitable for modeling environmental data.
- The models are introduced by means of integrals with respect to independently scattered stochastic measures that have generalized Laplace distributions.
- Resulting stationary second order processes have, as opposed to their Gaussian counterpart, a possibility of accounting for asymmetry and heavier tails.
- Despite this greater flexibility the discussed models still share a lot of spectral properties with Gaussian processes having the latter as a special case.
- The models extend directly to random fields.

Highlights

- Non-Gaussian stochastic fields are proposed that can be suitable for modeling environmental data.
- The models are introduced by means of integrals with respect to independently scattered stochastic measures that have generalized Laplace distributions.
- Resulting stationary second order processes have, as opposed to their Gaussian counterpart, a possibility of accounting for asymmetry and heavier tails.
- Despite this greater flexibility the discussed models still share a lot of spectral properties with Gaussian processes having the latter as a special case.
- The models extend directly to random fields.
- Spatio-temporal characteristics including asymmetries in the records can be studied by the means of generalized Rice's formula.

Highlights

- Non-Gaussian stochastic fields are proposed that can be suitable for modeling environmental data.
- The models are introduced by means of integrals with respect to independently scattered stochastic measures that have generalized Laplace distributions.
- Resulting stationary second order processes have, as opposed to their Gaussian counterpart, a possibility of accounting for asymmetry and heavier tails.
- Despite this greater flexibility the discussed models still share a lot of spectral properties with Gaussian processes having the latter as a special case.
- The models extend directly to random fields.
- Spatio-temporal characteristics including asymmetries in the records can be studied by the means of generalized Rice's formula.
- Model fitting can be obtained by utilizing: a) Method of Moments (simpler) or b) Maximum Likelihood (more accurate)

List of related papers

- Baxevani, A., Podgórski, K. and Wegener, J. (2014). Sample Path Asymmetries in Non-Gaussian Random Processes. *Scandinavian Journal of Statistics*
- Kotz, S., Kozubowski, T.J., Podgórski, K. (2001). *The Laplace distribution and generalizations: A revisit with applications to communications, economics, engineering and finance.* Boston: Birkhaüser.
- Åberg, S. and Podgórski, K. (2011). A class of non-gaussian second order random fields. *Extremes*
- Podgórski, K. and Wegener, J. (2010). Estimation for stochastic models driven by Laplace motion. *Comm. Statist. - Theory and Methods*.

List of related papers

- Baxevani, A., Podgórski, K. and Wegener, J. (2014). Sample Path Asymmetries in Non-Gaussian Random Processes. *Scandinavian Journal of Statistics*
- Kotz, S., Kozubowski, T.J., Podgórski, K. (2001). *The Laplace distribution and generalizations: A revisit with applications to communications, economics, engineering and finance.* Boston: Birkhaüser.
- Åberg, S. and Podgórski, K. (2011). A class of non-gaussian second order random fields. *Extremes*
- Podgórski, K. and Wegener, J. (2010). Estimation for stochastic models driven by Laplace motion. *Comm. Statist. - Theory and Methods*.
  Lindgren, F., Rue, H., and Lindström, J. (2011). An explicit link between gaussian fields and gaussian markov random fields: the stochastic partial differential equation approach. *J. Royal Stat. Soc., Series B.*
  Lagrange models and nested SPDE models. *European Phys. J.*

List of related papers

- Baxevani, A., Podgórski, K. and Wegener, J. (2014). Sample Path Asymmetries in Non-Gaussian Random Processes. *Scandinavian Journal of Statistics*
- Kotz, S., Kozubowski, T.J., Podgórski, K. (2001). *The Laplace distribution and generalizations: A revisit with applications to communications, economics, engineering and finance.* Boston: Birkhaüser.
- Åberg, S. and Podgórski, K. (2011). A class of non-gaussian second order random fields. *Extremes*
- Podgórski, K. and Wegener, J. (2010). Estimation for stochastic models driven by Laplace motion. *Comm. Statist. - Theory and Methods*.
  Lindgren, F., Rue, H., and Lindström, J. (2011). An explicit link between gaussian fields and gaussian markov random fields: the stochastic partial differential equation approach. *J. Royal Stat. Soc., Series B.*
  Lagrange models and nested SPDE models. *European Phys. J.*
- Bolin, D. (2011). Non-gaussian spatial models generated by stochastic partial differential equations.
- Podgórski, K., Walin, J. (2011). The EM-algorithm for models involving generalized Laplace distribution. *In progress*

List of related papers

- Baxevani, A., Podgórski, K. and Wegener, J. (2014). Sample Path Asymmetries in Non-Gaussian Random Processes. *Scandinavian Journal of Statistics*
- Kotz, S., Kozubowski, T.J., Podgórski, K. (2001). *The Laplace distribution and generalizations: A revisit with applications to communications, economics, engineering and finance.* Boston: Birkhaüser.
- Åberg, S. and Podgórski, K. (2011). A class of non-gaussian second order random fields. *Extremes*
- Podgórski, K. and Wegener, J. (2010). Estimation for stochastic models driven by Laplace motion. *Comm. Statist. - Theory and Methods*.
  Lindgren, F., Rue, H., and Lindström, J. (2011). An explicit link between gaussian fields and gaussian markov random fields: the stochastic partial differential equation approach. *J. Royal Stat. Soc., Series B.*
  Lagrange models and nested SPDE models. *European Phys. J.*
- Bolin, D. (2011). Non-gaussian spatial models generated by stochastic partial differential equations.
- Podgórski, K., Walin, J. (2011). The EM-algorithm for models involving generalized Laplace distribution. *In progress*
- Baxevani A. and Podgórski, K., (2014). Spectral representation of Laplace moving average process. *In progress*