Applications of the Rényi divergences in testing hypotheses about exponential models

Třešť, December 12, 2006

Lucie Fajfrová
ÚTIA AV ČR

1. Rényi divergences

2.

3.

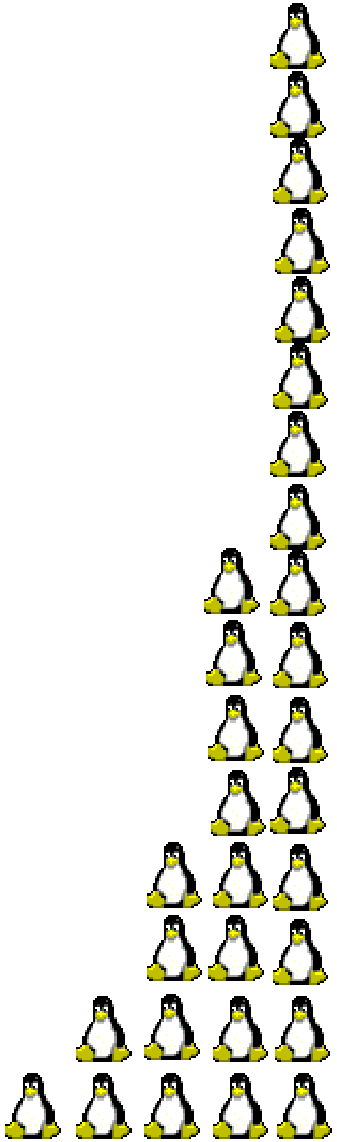
4.

5.

Alfréd Rényi

1. Rényi divergences

2. Exponential families



3.

4.

5.

1. Rényi divergences

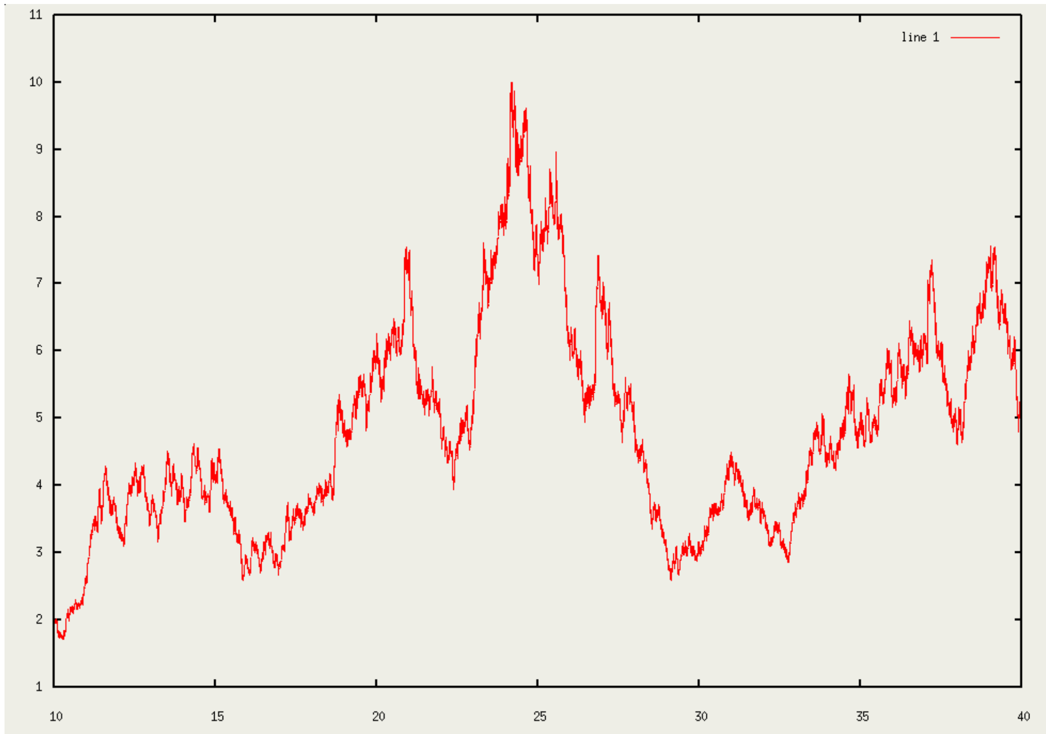
2. Exponential families

   - Exponential families of processes

3.

4.

5.



1. Rényi divergences

2. Exponential families

   - Exponential families of processes

   - Lévy processes

3.

4.

5.

Paul Lévy

1. Rényi divergences

2. Exponential families

    - Exponential families of processes

    - Lévy processes

3. Rényi statistics

4. Calculation

5. Another examples

Rényi divergences

of order $r \in \mathbb{R}$ for probability measures P, P_0 with densities f, f_0 on a σ-finite measure space $(\mathfrak{X}, \mathcal{A}, \mu)$ is:

$$D_r(P, P_0) = \frac{1}{r(r-1)} \ln \int f^r f_0^{1-r} \, d\mu,$$

for $r \neq 1, r \neq 0$,

$$D_1(P, P_0) = \int f \ln \frac{f}{f_0} \, d\mu$$

and

$$D_0(P, P_0) = \int f_0 \ln \frac{f_0}{f} \, d\mu.$$

A. Rényi. On measures of entropy and information.1961

F. Liese, I. Vajda. *Convex statistical distances.* 1987.

Exponential families

$\{P_\theta : \theta \in \Theta\}$ on $(\mathfrak{X}, \mathcal{A}, \mu)$

•

–

–

•

–

–

–

Exponential families

$\{P_\theta : \theta \in \Theta\}$ on $(\mathfrak{X}, \mathcal{A}, \mu)$

$$\frac{dP_\theta}{d\mu} =: f_\theta(x)$$

-
  -
  -
-
  -
  -
  -

Exponential families

$$\{P_\theta : \theta \in \Theta\} \quad \text{on} \quad (\mathfrak{X}, \mathcal{A}, \mu)$$

iff

$$\frac{dP_\theta}{d\mu} =: f_\theta(x) = \exp\{\theta' T(x) - \kappa(\theta)\} \text{ for each } \theta \in \Theta$$

•

–

–

•

–

–

–

Exponential families

$$\{P_\theta : \theta \in \Theta\} \quad \text{on} \quad (\mathfrak{X}, \mathcal{A}, \mu)$$

iff

$$\frac{dP_\theta}{d\mu} =: f_\theta(x) = \exp\{\theta' T(x) - \kappa(\theta)\} \text{ for each } \theta \in \Theta$$

- models with iid observations
  - normal, lognormal, gamma, beta, Weibull, Maxwell,...
  - Poisson, negative binomial, multinomial,...
- models with dependent observations
  - sequences of dependent observations
  - time continuous processes
  - random fields

Exponential families

$$\{P_\theta : \theta \in \Theta\} \quad \text{on} \quad (\mathfrak{X}, \mathcal{A}, \mu)$$

iff

$$\frac{dP_\theta}{d\mu} =: f_\theta(x) = \exp\{\theta' T(x) - \kappa(\theta)\} \text{ for each } \theta \in \Theta$$

- models with iid observations
  - normal, lognormal, gamma, beta, Weibull, Maxwell,...
  - Poisson, negative binomial, multinomial,...
- models with dependent observations
  - sequences of dependent observations
  - time continuous processes
  - random fields

Exponential families of processes

$\{P_\theta : \theta \in \Theta\}$ on filtered canonical path-space $(\mathfrak{X}, \mathcal{A}, (\mathcal{A}_t)_{t\geq 0}, \mu)$

iff

$$\frac{dP_{\theta,t}}{d\mu_t} =: f_{\theta,t}(\mathbb{X}_t) = \exp\{\theta' T_t(\mathbb{X}_t) - \kappa(\theta) S_t(\mathbb{X}_t)\}$$

for each $\theta \in \Theta$ for each $t > 0$

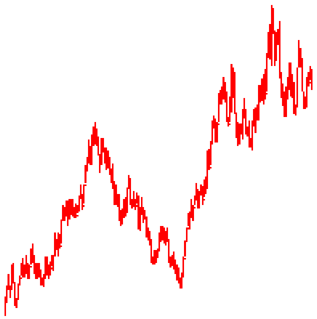
Exponential families of processes

$\{P_\theta : \theta \in \Theta\}$ on filtered canonical path-space $(\mathfrak{X}, \mathcal{A}, (\mathcal{A}_t)_{t\geq 0}, \mu)$

iff

$$\frac{dP_{\theta,t}}{d\mu_t} =: f_{\theta,t}(\mathbb{X}_t) = \exp\{\theta' T_t(\mathbb{X}_t) - \kappa(\theta) S_t(\mathbb{X}_t)\}$$

for each $\theta \in \Theta$ for each $t > 0$



$\mathbb{X}_t$ is observation up to time t

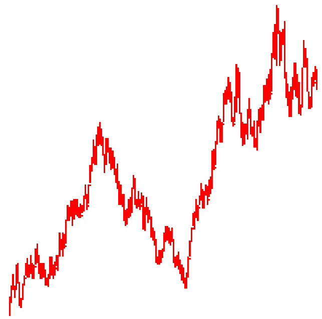
Exponential families of processes

$\{P_\theta : \theta \in \Theta\}$ on filtered canonical path-space $(\mathfrak{X}, \mathcal{A}, (\mathcal{A}_t)_{t\geq 0}, \mu)$

iff

$$\frac{dP_{\theta,t}}{d\mu_t} =: f_{\theta,t}(\mathbb{X}_t) = \exp\{\theta' T_t(\mathbb{X}_t) - \kappa(\theta) S_t(\mathbb{X}_t)\}$$

for each $\theta \in \Theta$ for each $t > 0$



$\mathbb{X}_t$ is observation up to time t

$S_t \geq 0$
- natural models: $S_t = S_t(\mathbb{X}_t) = t$ — Lévy processes
- general models: S_t is nondecreasing (time transformation)

Natural models, Lévy processes

$$\frac{dP_{\theta,t}}{dP_{0,t}} = \exp\{\theta' T_t(\mathbb{X}_t) - \kappa(\theta)t\}$$

Lévy processes ...... $T_t(\mathbb{X}_t) = X_t$

- Brownian motion with unknown drift
- Poisson process with unknown intensity of jumps
- their combinations

Natural models, Lévy processes ($T_t(\mathbb{X}_t) = X_t$)

$$\frac{dP_{\theta,t}}{dP_{0,t}} = \exp\{\theta' T_t(\mathbb{X}_t) - \kappa(\theta)t\}$$

Rényi divergences in natural models

For all real $r \neq 0$, $r \neq 1$

$$D_r(P_{\theta_1}, P_{\theta_0}) := D_r(\theta_1, \theta_0) = \frac{\kappa(\, r\theta_1 + (1-r)\theta_0\,) - r\kappa(\theta_1) - (1-r)\kappa(\theta_0)}{r(r-1)}$$

if $r\theta + (1-r)\theta_0 \in \Theta$,
otherwise $D_r(\theta_1, \theta_0) = \infty$.

$$D_1(\theta_1, \theta_0) = \dot{\kappa}(\theta_1)\,(\theta_1 - \theta_0) + \kappa(\theta_0) - \kappa(\theta_1)$$

Usual statistical features

- likelihood function for natural exponential models:

$$L_t(\theta) = \exp\{\theta' T_t - \kappa(\theta) t\}$$

- maximum likelihood estimator:

$$\widehat{\theta}_t = \underset{\theta \in \Theta}{\operatorname{argmax}} \; L_t(\theta) \; = \dot{\kappa}^{-1}\left(\frac{T_t}{t}\right)$$

- testing hypotheses $\quad H_0 : \quad g(\theta) = 0$

where g is continuous, 2x diff. function $\mathbb{R}^k \to \mathbb{R}^{k-\tilde{k}}$, $1 \le \tilde{k} < k$.

- generalised likelihood estimator:

$$\tilde{\theta}_t = \underset{g(\theta)=0}{\operatorname{argmax}} \; L_t(\theta)$$

Rényi statistics

- for testing simple hypotheses $H_0 : \theta = \theta_0$

Rényi statistics of orders $r \in \mathbb{R}$: $D_{r,t} = 2 * t * D_r(\widehat{\theta}_t, \theta_0)$,

Kullbac statistics: $K_t = -2 * \log \left(\frac{L_t(\theta_0)}{L_t(\widehat{\theta}_t)} \right) = 2 * t * D_1(\widehat{\theta}_t, \theta_0)$

Asymptotic behaviour ($t \to \infty$): $D_{r,t} \xrightarrow{\mathcal{L}} \chi_1^2$

(under some regularity conditions)

.

Rényi statistics

- • for testing composite hypotheses $H_0 : g(\theta) = 0$

.

Rényi statistics

• • for testing composite hypotheses $H_0 : g(\theta) = 0$

Rényi statistics of orders $r \in \mathbb{R}$: $D_{r,t} = 2 * t * D_r(\widehat{\theta}_t, \widetilde{\theta}_t)$,

Kullbac statistics: $K_t = -2 * \log\left(\frac{L_t(\widetilde{\theta}_t)}{L_t(\widehat{\theta}_t)}\right) = 2 * t * D_1(\widehat{\theta}_t, \widetilde{\theta}_t)$

Asymptotic behaviour ($t \to \infty$): $D_{r,t} \xrightarrow{\mathcal{L}} \chi^2_{k-\tilde{k}}$

(under some regularity conditions)

.

Rényi statistics

• • for testing composite hypotheses $H_0 : g(\theta) = 0$

Rényi statistics of orders $r \in \mathbb{R}$: $D_{r,t} = 2 * t * D_r(\widehat{\theta}_t, \tilde{\theta}_t)$,

Kullbac statistics: $K_t = -2 * \log \left(\frac{L_t(\tilde{\theta}_t)}{L_t(\widehat{\theta}_t)} \right) = 2 * t * D_1(\widehat{\theta}_t, \tilde{\theta}_t)$

Asymptotic behaviour ($t \to \infty$): $D_{r,t} \xrightarrow{\mathcal{L}} \chi^2_{k-\tilde{k}}$

(under some regularity conditions)

U. Küchler, M. Sørensen. Exponential Families of Stochastic Processes. 1997.

Rényi statistics

- • for testing composite hypotheses $H_0 : g(\theta) = 0$

Rényi statistics of orders $r \in \mathbb{R}$: $D_{r,t} = 2 * t * D_r(\widehat{\theta}_t, \widetilde{\theta}_t)$,

Kullbac statistics: $$K_t = -2 * \log\left(\frac{L_t(\widetilde{\theta}_t)}{L_t(\widehat{\theta}_t)}\right) = 2 * t * D_1(\widehat{\theta}_t, \widetilde{\theta}_t)$$

Asymptotic behaviour ($t \to \infty$):

(under some regularity conditions)

$$D_{r,t} \xrightarrow{\mathcal{L}} \chi^2_{k-\tilde{k}}$$

D. Morales, L. Pardo, I. Vajda. Some new statistics for testing ... 1997.

Rényi statistics

- • for testing composite hypotheses $H_0 : g(\theta) = 0$

Rényi statistics of orders $r \in \mathbb{R}$: $D_{r,t} = 2 * t * D_r(\widehat{\theta}_t, \widetilde{\theta}_t)$,

Kullbac statistics: $K_t = -2 * \log \left(\dfrac{L_t(\widetilde{\theta}_t)}{L_t(\widehat{\theta}_t)} \right) = 2 * t * D_1(\widehat{\theta}_t, \widetilde{\theta}_t)$

Asymptotic behaviour ($t \to \infty$): $D_{r,t} \xrightarrow{\mathcal{L}} \chi^2_{k-\tilde{k}}$

(under some regularity conditions)

D. Morales, L. Pardo, I. Vajda. Rényi statistics in directed families ... 2000.

Rényi statistics

- • for testing composite hypotheses $H_0 : g(\theta) = 0$

Rényi statistics of orders $r \in \mathbb{R}$: $D_{r,t} = 2 * t * D_r(\widehat{\theta}_t, \widetilde{\theta}_t)$,

Kullbac statistics: $K_t = -2 * \log \left(\frac{L_t(\widetilde{\theta}_t)}{L_t(\widehat{\theta}_t)} \right) = 2 * t * D_1(\widehat{\theta}_t, \widetilde{\theta}_t)$

Asymptotic behaviour ($t \to \infty$): $D_{r,t} \xrightarrow{\mathcal{L}} \chi^2_{k-\tilde{k}}$

(under some regularity conditions)

D. Morales, L. Pardo, M.C. Pardo, I. Vajda. Rényi statistics for testing ... 2004.

Main example

Geometrical Brownian motion with Poisson jumps

- for modelling some phenomena in financial markets
- as a model for the stock price
- fitting on evolution of petroleum price

Main example

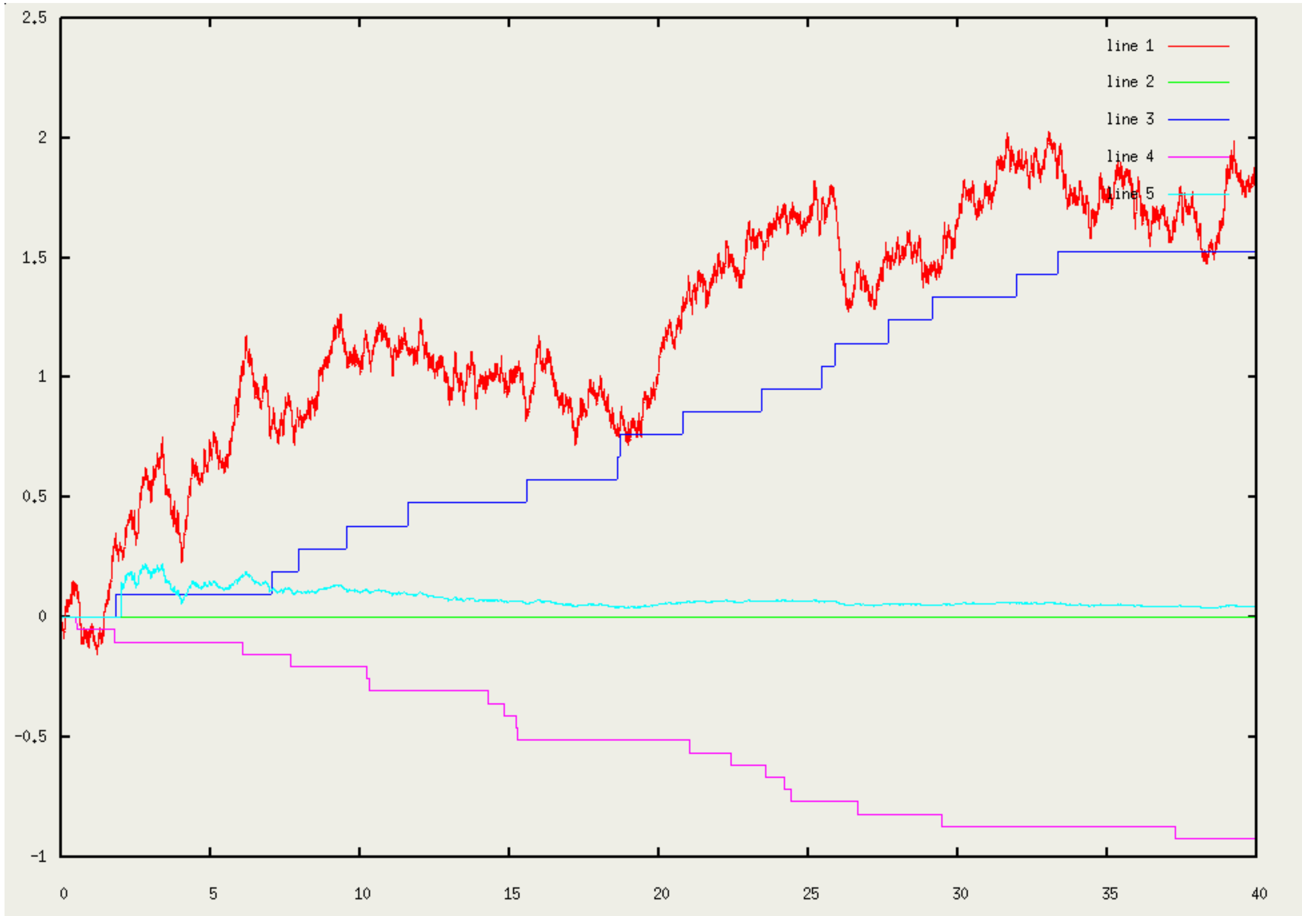
Geometrical Brownian motion with Poisson jumps

- for modelling some phenomena in financial markets
- as a model for the stock price
- fitting on evolution of petroleum price

It has three components : drifted Brownian motion
Poisson process of jumps up
Poisson process of jumps down

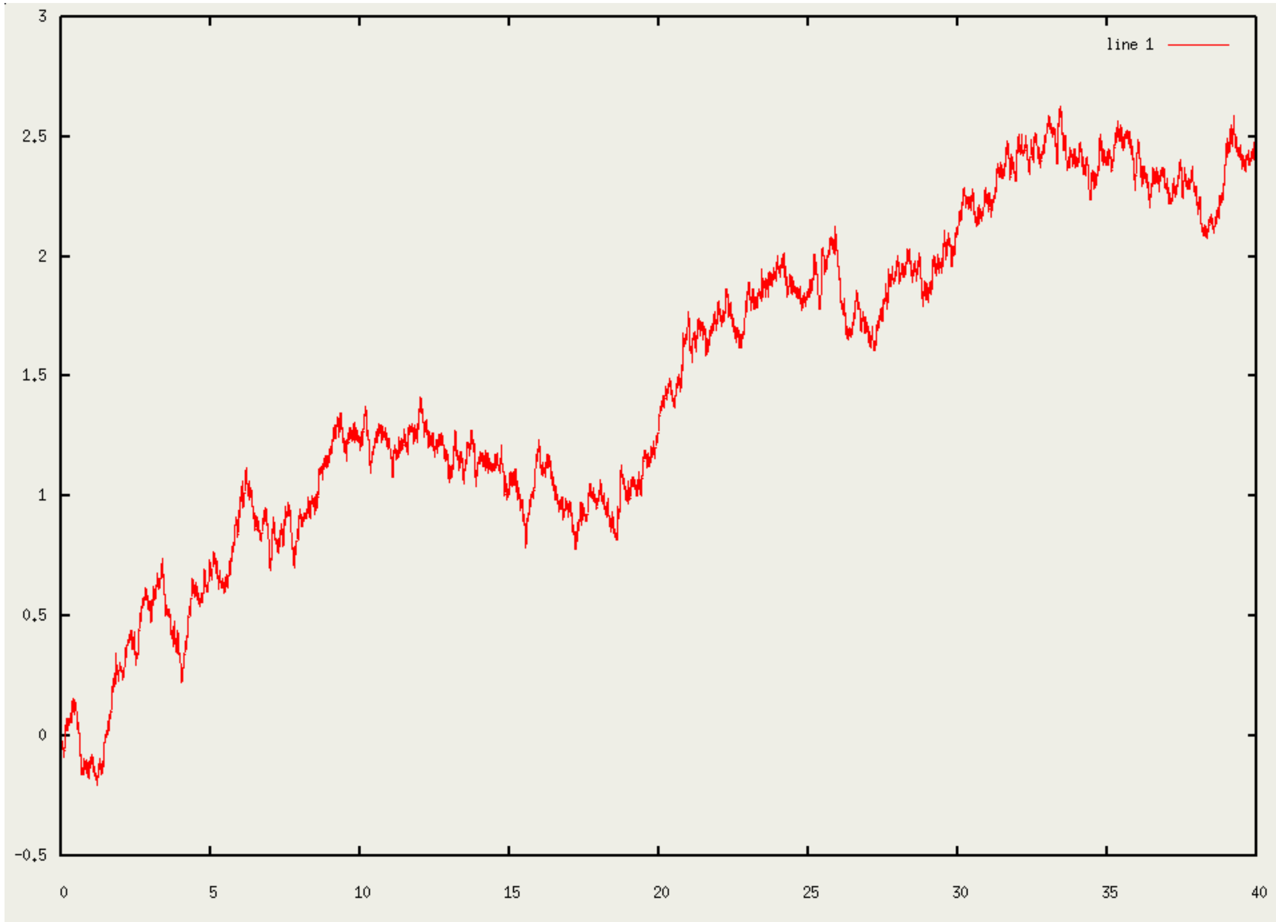
Geometrical Brownian motion with Poisson jumps

drift$= 0.05$, $\lambda_{up} = 0.72$, $\lambda_{down} = 0.48$



Geometrical Brownian motion with Poisson jumps

drift$= 0.05$, $\lambda_{up} = 0.72$, $\lambda_{down} = 0.48$



Geometrical Brownian motion with Poisson jumps

Description of the model

X_t is solution of SDE

$$\begin{aligned} dX_t &= mX_t dt + \sigma X_t dW_t + 0.1 X_t dN_t^{\lambda p} - 0.05 X_t dN_t^{\lambda q} \\ X_0 &= x_0 \end{aligned}$$

m … proportional drift, σ … proportional diffusion, $p + q = 1$

Explicitly

$$X_t = x_0 \exp\Big\{ \underbrace{(m - \sigma^2/2)}_{=:\theta} t + \sigma W_t + \log(1.1)\, N_t^{\lambda p} + \log(0.95)\, N_t^{\lambda q} \Big\}$$

Unknown parameters:

- ▷ new drift θ
- ▷ intensity of jumps λ
- ▷ probability of jumping up p

Transformation

- ∗ $\theta_1 = \theta$
- ∗ $\theta_2 = \log(\lambda p)$
- ∗ $\theta_3 = \log(\lambda q)$

Geometrical Brownian motion with Poisson jumps

Statistics

data: $N_t^{\lambda p}$ & $N_t^{\lambda q}$ & $B_t = \theta t + \sigma W_t$

MLE: $$(\widehat{\theta}_{1,t}, \widehat{\theta}_{2,t}, \widehat{\theta}_{3,t}) = \left(\frac{B_t}{t},\ \log(\frac{N_t^{\lambda p}}{t}),\ \log(\frac{N_t^{\lambda q}}{t}) \right)$$

hypothesis: $H_0 :\ \theta = \lambda p \log(1.1) + \lambda q \log(0.95) \quad \& \quad p = 2/3$

$$H_0 :\ \theta_1 = e^{\theta_2} \log(1.1) + e^{\theta_3} \log(0.95) \quad \& \quad e^{\theta_2} = 2e^{\theta_3}$$

Geometrical Brownian motion with Poisson jumps

Statistics

data: $N_t^{\lambda p}$ & $N_t^{\lambda q}$ & $B_t = \theta t + \sigma W_t$

MLE: $$(\widehat{\theta}_{1,t}, \widehat{\theta}_{2,t}, \widehat{\theta}_{3,t}) = \left(\frac{B_t}{t},\ \log(\frac{N_t^{\lambda p}}{t}),\ \log(\frac{N_t^{\lambda q}}{t})\right)$$

hypothesis: $H_0 :\ \theta = \lambda p \log(1.1) + \lambda q \log(0.95)$ & $p = 2/3$

$$H_0 :\ \theta_1 = e^{\theta_2} \log(1.1) + e^{\theta_3} \log(0.95) \quad \& \quad e^{\theta_2} = 2e^{\theta_3}$$

Geometrical Brownian motion with Poisson jumps

Statistics

data: $N_t^{\lambda p}$ & $N_t^{\lambda q}$ & $B_t = \theta t + \sigma W_t$

MLE: $(\widehat{\theta}_{1,t}, \widehat{\theta}_{2,t}, \widehat{\theta}_{3,t}) = \left(\frac{B_t}{t},\ \log(\frac{N_t^{\lambda p}}{t}),\ \log(\frac{N_t^{\lambda q}}{t}) \right)$

hypothesis: $H_0: \ \theta = \lambda p \log(1.1) + \lambda q \log(0.95) \quad \& \quad p = 2/3$

$$H_0: \ \theta_1 = e^{\theta_2} \log(1.1) + e^{\theta_3} \log(0.95) \quad \& \quad e^{\theta_2} = 2e^{\theta_3}$$

Geometrical Brownian motion with Poisson jumps

Statistics

data: $N_t^{\lambda p}$ & $N_t^{\lambda q}$ & $B_t = \theta t + \sigma W_t$

MLE: $$(\widehat{\theta}_{1,t}, \widehat{\theta}_{2,t}, \widehat{\theta}_{3,t}) = \left(\frac{B_t}{t},\ \log(\frac{N_t^{\lambda p}}{t}),\ \log(\frac{N_t^{\lambda q}}{t})\right)$$

hypothesis: $H_0:\ \theta = \lambda p \log(1.1) + \lambda q \log(0.95) \quad \& \quad p = 2/3$

$$H_0:\ \theta_1 = e^{\theta_2} \log(1.1) + e^{\theta_3} \log(0.95) \quad \& \quad e^{\theta_2} = 2e^{\theta_3}$$

Geometrical Brownian motion with Poisson jumps

Statistics

Rényi statistics: $$D_{r,t} = \frac{2t}{r(r-1)}[\kappa(\underline{\theta}_r) - r\kappa(\underline{\widehat{\theta}}_t) - (1-r)\kappa(\underline{\widetilde{\theta}}_t)]$$

$r \neq 0, 1$

Kullback statistics: $$D_{1,t} = 2t\left[\dot{\kappa}(\underline{\widehat{\theta}}_t)(\underline{\widehat{\theta}}_t - \underline{\widetilde{\theta}}_t) + \kappa(\underline{\widetilde{\theta}}_t) - \kappa(\underline{\widehat{\theta}}_t)\right]$$

where $$\kappa(\underline{\theta}) = \kappa(\theta_1, \theta_2, \theta_3) = \frac{\theta_1^2}{2\sigma^2} + e^{\theta_2} + e^{\theta_3} - 2$$

$$\underline{\theta}_r = r\underline{\widehat{\theta}}_t + (1-r)\underline{\widetilde{\theta}}_t$$

$$\underline{\widetilde{\theta}}_t = (\widetilde{\theta}_{1,t}, \widetilde{\theta}_{2,t}, \widetilde{\theta}_{3,t}) = \underset{\substack{e^{\theta_2} = 2e^{\theta_3} \\ \theta_1 = e^{\theta_2}\log(1.1) + e^{\theta_3}\log(0.95)}}{\operatorname{argmax}} \left(\theta_1 \frac{B_t}{\sigma^2} + \theta_2 N_t^{\lambda p} + \theta_3 N_t^{\lambda q} - t\left(\kappa(\theta_1, \theta_2, \theta_3)\right)\right)$$

Geometrical Brownian motion with Poisson jumps

Calculations

• simulating of independent realisations from normalized normal and Poisson distributions

• calculation of test powers for observations up to time $t = 40$ for Rényi statistics of orders $r \in \{-1;\ 0.5;\ 1;\ 1.5;\ 2;\ 2.5;\ \}$

• optimization: comparison between test powers

$$\Downarrow$$

choice of statistics $D_{0.5,40}$

Order -1	0,5	1	1,5	2	2,5		p	Intensity	Drift
~ of Renyi									
0,9948	0,9954	0,9956	0,9958	0,9959	0,9961		0,45	3,1	0,16
0,9690	0,9721	0,9733	0,9748	0,9756	0,9765		0,45	2,4	0,11
0,8990	0,9103	0,9143	0,9189	0,9235	0,9280		0,45	1,7	0,06
0,7975	0,8068	0,8124	0,8195	0,8261	0,8335		0,45	1,3	0,06
0,6650	0,6768	0,6824	0,6876	0,6931	0,6988		0,55	3,1	0,16
0,5278	0,5431	0,5520	0,5593	0,5704	0,5818		0,55	2,4	0,11
0,4114	0,4378	0,4474	0,4575	0,4710	0,4831		0,55	1,7	0,06
0,3061	0,3204	0,3286	0,3380	0,3486	0,3599		0,55	1,3	0,06
0,0818	0,0791	0,0800	0,0806	0,0810	0,0814		0,67	1,7	0,06
0,0558	0,0506	0,0503	0,0503	0,0501	0,0493		0,67	2,4	0,11
0,0574	0,0571	0,0571	0,0570	0,0575	0,0580		0,67	2,6	0,11
0,0768	0,0693	0,0678	0,0660	0,0649	0,0641		0,67	3,1	0,16
0,3303	0,2911	0,2786	0,2673	0,2566	0,2459		0,75	1,7	0,06
0,3849	0,3415	0,3270	0,3135	0,2990	0,2871		0,75	2,4	0,11
0,4204	0,3830	0,3699	0,3590	0,3488	0,3388		0,75	2,6	0,11
0,5200	0,4733	0,4598	0,4455	0,4304	0,4151		0,75	3,1	0,16
0,8680	0,8188	0,8029	0,7834	0,7664	0,7446		0,85	1,3	0,06
0,9326	0,9103	0,9018	0,8935	0,8864	0,8774		0,85	1,7	0,06
0,9819	0,9745	0,9711	0,9671	0,9644	0,9593		0,85	2,4	0,11
0,9976	0,9963	0,9954	0,9948	0,9944	0,9936		0,85	3,1	0,16

Some more examples

- on behalf of diffusion process

Ornstein-Uhlenbeck process

is the unique solution of SDE
$$\begin{aligned} dX_t &= \theta X_t dt + dW_t \\ X_0 &= 0 \end{aligned}$$
$\theta < 0$

Explicitly $X_t = \int\limits_0^t \exp(\theta(t-s))dW_s.$

Density and likelihood function

$$\frac{dP_t^\theta}{dP_t^0}(\mathbb{X}_t) = \exp\left(\theta\frac{1}{2}(X_t^2 - t) - \frac{1}{2}\theta^2\int\limits_0^t X_s^2 ds\right)$$

MLE

$$\hat{\theta} = \frac{X_t^2 - t}{2\int\limits_0^t X_s^2 ds}$$

Some more examples

- on behalf of counting processes

Pure birth process

Let $X_0 = 1$ which means that there is a population with just one individual at time $t = 0$.

$X_t \in \{1, 2, ...\}$ then means the size of population at time t, if each individual gives birth to another individual after an exponential time with expectation $1/\lambda$. Behaviour of individuals is independent. $\rightsquigarrow \theta := \log \lambda \in \mathbb{R}$

Density and likelihood function

$$\frac{dP_t^\theta}{dP_t^0}(\mathbb{X}_t) = \exp\{\theta(X_t - 1) - (e^\theta - 1) \int_0^t X_s ds\}$$

MLE

$$\widehat{\theta} = \log \frac{X_t - 1}{\int_0^t X_s ds}$$

Future plans

∗ Could be used Rényi divergences for testing hypotheses also in general exponential models?

∗ Is there again some nice formula for the test statistics?

∗ Asymptotics for them?

∗ How to handle with that random time transformation S_t?

• Which models are the most suitable for using Rényi statistics?

U. Küchler, M. Sørensen. *Exponential Families of Stochastic Processes*. Springer, Berlin, 1997.

F. Liese, I. Vajda. *Convex statistical distances.* Teubner, Leipzig, 1987.

D. Morales, L. Pardo, I. Vajda. Some new statistics for testing hypothesis in parametric models. *Journal of Multivariate Analysis*, 62: 137 – 168, 1997.

D. Morales, L. Pardo, I. Vajda. Rényi statistics in directed families of exponential experiments. *Statistics*, 34: 151 – 174, 2000.

D. Morales, L. Pardo, M.C. Pardo, I. Vajda. Rényi statistics for testing composite hypothesis in general exponential models. *Statistics*, 38, No.2: 133 – 147, 2004.

A. Rényi. On measures of entropy and information. *Proc. 4th Berkeley Symp.Math.Statist.Probab.*, 1: 547 – 561, 1961.