“A practical approach to the inference for jump-diffusions from finite samples ”

Yasutaka SHIMIZU*
Graduate School of Engineering Science,
Osaka University
yasutaka@sigmath.es.osaka-u.ac.jp

Asymptotical Statistics of Stochastic Processes VI
Universite du Maine, Le Mans, 21-23 March, 2007

Introduction:

- Recently, modeling by **jump-processes** are frequently used in finance and insurance, and the statistical infenrence for such processes is important.
- In applications, the observational data are usually obtained discretely, so we need an estimation procedure for jump-processes from **discrete observations**.
- Recently, Shimizu and Yoshida (2006) studied the **asymptotic theory** for discretely observed jump-diffusions, but, in practice, the performance of their estimation method is sometimes not so good since the real data are finite samples.
- We would like to put their theory to practical use, and to implement the 'method' in this talk as a **computer algorithm**.

1. Problem

Inference for Jump-Diffusions

We deal with 1-dim SDE on $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\geq 0}, P)$:

$$dX_t = a(X_t)\,dt + b(X_t)\,dw_t + dz_t,$$

where:

· a, b: **unknown** functions. · w: Wiener process.

· $z_t = \sum_{i=1}^{N_t} \varepsilon_i$: Compound Poisson process;

N_t: $\sim Po(\lambda t)$, ε_i: i.i.d. $\sim F(z)\,dz$.

$$f(z) := \lambda F(z) \ : \ \textbf{unknown} \text{ Lévy density.}$$

Discrete observations

$$\boldsymbol{X}^n = \left\{ X_{t_i^n} \right\}_{i=0,1,\cdots n}, \qquad t_i^n = ih_n \ \ (h_n > 0).$$

Consider the asymptotics that

$$h_n \to 0, \qquad nh_n \to \infty \qquad (n \to \infty).$$

<u>Recent literatures</u>

- via Local time :Bandi and Nguyen (2003)
- λ-estimation : Mancini (2004)
- MLE-type : Shimizu and Yoshida (2006)
- Density estimation : Shimizu (2006) etc.

Judgement of jumps

$r_n(\vartheta)$: A sequence (ϑ :A parameter),
$\Delta_i X^n := X_{t_i^n} - X_{t_{i-1}^n}$: Increment in $[t_{i-1}^n, t_i^n)$.

Asymptotic filter: $n \to \infty$

$$|\Delta_i X^n| \le r_n(\vartheta) \Rightarrow \text{ No jump in } (t_{i-1}^n, t_i^n],$$
$$|\Delta_i X^n| > r_n(\vartheta) \Rightarrow \text{ A single jump in } (t_{i-1}^n, t_i^n],$$

and $\Delta_i X^n \approx$ jump size.

Mancini (2004):
For any $\vartheta > \exists c$, $r_n(\vartheta) = \vartheta\sqrt{h_n}\log h_n^{-1}$.

Shimizu and Yoshida (2006):
For any $\vartheta_1 > 0$ and any $\vartheta_2 \in (0, 1/2)$, $r_n(\vartheta) = \vartheta_1 h_n^{\vartheta_2}$ $\vartheta = (\vartheta_1, \vartheta_2)$.

Roughly speaking, when $|\Delta_i X^n| > r_n(\vartheta)$,

$$\Delta_i X^n \approx \text{Jump size} \quad \Rightarrow \quad \lim_{n\to\infty} r_n(\vartheta) = \lim_{n\to\infty} \frac{\sqrt{h_n}}{r_n(\vartheta)} = 0.$$

Q. How to select ϑ?

· Asymptotically (theoretically), **any** ϑ is OK!

· In practice, we have to detrmin ϑ suitably under **fixed** n.

$\Rightarrow$ See the next example.

2. Example (SY,2006)

Consider

$$dX_t = \mu X_t \, dt + \sigma \, dW_t + dz_t^\nu,$$

where z^ν is a compound Poisson process with Lévy density

$$f_\nu(z) = \frac{\lambda}{\sqrt{2\pi\nu_2}} \exp\left(-\frac{(z-\nu_1)^2}{2\nu_2} \right).$$

$h_n = n^{-0.7}$ $(nh_n^2 \to 0)$, and $r_n(\vartheta) = \vartheta h_n^{0.49}$.

Experiments: 500 times.

	n (nh_n)	50 (3.2)	500 (6.5)	3000 (11.0)	True
σ =0.1	$\hat{\mu}_n$	-0.2702	-0.3009	-0.3013	-0.3
ϑ =1.0	s.d.	0.1474	0.0477	0.0277	
	$\hat{\sigma}_n$	0.1169	0.1003	0.0998	0.1
	s.d.	0.0295	0.0066	0.0025	
	$\hat{\nu}_{1,n}$	0.6450	0.5408	0.5135	0.5
	s.d.	0.1294	0.0745	0.0558	
	$\hat{\nu}_{2,n}$	0.0919	0.0924	0.0956	0.1
	s.d.	0.0886	0.0495	0.0269	
	$\hat{\lambda}_n$	2.0659	2.6893	2.9067	3.0
	s.d.	0.7266	0.6514	0.5076	

	n (nh_n)	50 (3.2)	500 (6.5)	3000 (11.0)	True
σ =0.3	$\hat{\mu}_n$	-0.2358	-0.2528	-0.2536	-0.3
ϑ =1.0	s.d.	0.2366	0.0809	0.0428	
	$\hat{\sigma}_n$	0.2310	0.2381	0.2471	0.3
	s.d.	0.0433	0.0133	0.0054	
	$\hat{\nu}_{1,n}$	0.4505	0.2230	0.1108	0.5
	s.d.	0.0151	0.0522	0.0162	
	$\hat{\nu}_{2,n}$	0.2131	0.1445	0.0839	0.1
	s.d.	0.1178	0.0331	0.0162	
	$\hat{\lambda}_n$	2.7419	5.7280	11.453	3.0
	s.d.	0.8759	0.9882	1.0682	

	n (nh_n)	50 (3.2)	500 (6.5)	3000 (11.0)	True
σ =0.3	$\hat{\mu}_n$	-0.2001	-0.2977	-0.3010	-0.3
ϑ =1.8	s.d.	0.2366	0.0866	0.0428	
	$\hat{\sigma}_n$	0.3902	0.3044	0.2978	0.3
	s.d.	0.0901	0.0191	0.0074	
	$\hat{\nu}_{1,n}$	0.7517	0.5750	0.5147	0.5
	s.d.	0.1879	0.0800	0.0552	
	$\hat{\nu}_{2,n}$	0.2014	0.0965	0.1002	0.1
	s.d.	0.2054	0.0531	0.0279	
	$\hat{\lambda}_n$	1.5156	2.4717	2.8938	3.0
	s.d.	0.6450	0.5881	0.4682	

<u>3. Selection of Filter</u>

A criterion to select the filters

According to the simulation results, we should select the constant ϑ suitably depending on the model.

Hearafter, we consider the following jump-judgment filter:

$$\mathcal{H}_i^n(r_n) = \{\omega \in \Omega\,;\, |\Delta_i X^n| > r_n\}\,.$$

We consider to choose a suitable threshold r_n for fixed n.

$\Rightarrow$ **How do we choose a suitable r_n?**
Asymptotic theory does not answer this question.

Here we notice the natural estimator of λ:

$$\hat{\lambda}(r_n) = \frac{1}{nh_n} \sum_{i=1}^{n} \mathbf{1}_{\mathcal{H}_i^n(r_n)},$$

This is an asymptotically efficient estimator (e.g. SY, 2006).

For fixed n, it is desired that the bias

$$b(r_n) := E\left[\hat{\lambda}(r_n)\right] - \lambda \tag{1}$$

is as small as possible.

Proposition 1 Let

$$\tilde{\ell}_n := h_n^{-1}\varepsilon(r_n) - \int_{|z|<2r_n} f(z)\,dz,$$
$$\ell_n := \tilde{\ell}_n + \int_{\frac{r_n}{2}<|z|\leq 2r_n} f(z)\,dz,$$

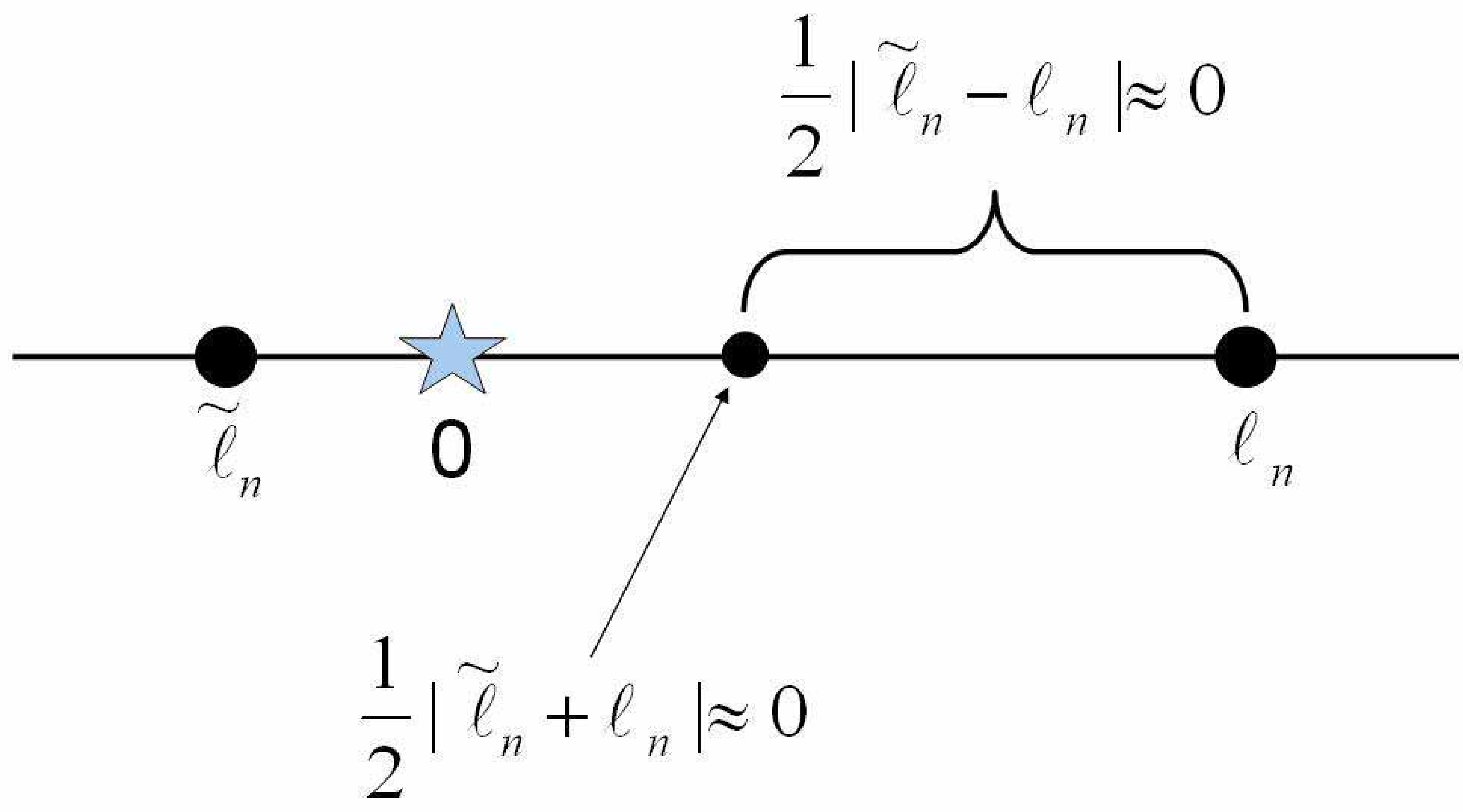
where

$$\varepsilon(r_n) = \frac{1}{n}\sum_{i=1}^{n} P\left\{|\Delta_i X^n| > r_n,\ \text{no jump in } [t_{i-1}^n, t_i^n)\right\}.$$

Then, as $h_n \to 0$,

$$\tilde{\ell}_n + O(h_n) \leq e^{\lambda h_n} b(r_n) \leq \ell_n + O(h_n), \tag{2}$$

- $\frac{1}{2}|\tilde{\ell}_n + \ell_n|$: Center of the bias' range.
- $\frac{1}{2}|\tilde{\ell}_n - \ell_n|$: Amplitude of the bias' range.



For $w \geq 0$, let

$$
\begin{aligned}
L_{w,n}[r_n, \varepsilon, f] &:= w|\tilde{\ell}_n + \ell_n| + (1-w)|\tilde{\ell}_n - \ell_n| \\
&= w\left| 2h_n^{-1}\varepsilon(r_n) - \int_{|z| \leq \frac{r_n}{2}} f(z)\,dz - \int_{|z|<2r_n} f(z)\,dz \right| \\
&\quad + (1-w)\int_{\frac{r_n}{2}<|z|\leq 2r_n} f(z)\,dz,
\end{aligned}
$$

$$
r_{opt}^{(n,w)} := \arg\min_{r \geq 0} L_{w,n}[r, \varepsilon, f],
$$

which is expected to be a *good* threshold.

Fix some $w \geq 0$ and find the r which mimimizes

$$L_{w,n}[r, \varepsilon, f] \qquad (r \geq 0).$$

$\Rightarrow$ Substitute ε and f with their consistent estimators:

$$L_{w,n}[r, \hat{\varepsilon}_n, \hat{f}_n].$$

However estimators $\hat{\varepsilon}_n$ and $\hat{f}_n$ are also to be constructed by using Filter like $\{|\Delta_i X^n| \leq r\}$ or $\{|\Delta_i X^n| > r\}$; **We want this $\boldsymbol{r}$!**.

$\Rightarrow$ It goes back and forth!

4. Plug-in Method

Notation: For an unknown function $g(x)$, denote by $\hat{g}_n(x; s)$ the estimator of $g(x)$ constructed via $\{|\Delta_i X^n| > s\}$ or $\{|\Delta_i X^n| \leq s\}$.

Direct plug-in rule

Step 0: Choose a constant $r_n^{(0)}$; **pilot threshold**, and make estimators $\hat{\varepsilon}_n(r; r_n^{(0)})$, $\hat{f}_n(z; r_n^{(0)})$. □

Step $\boldsymbol{k}$ ($k \geq 1$): Find the following $r_n^{(k)}$; **$\boldsymbol{k}$-stage threshold**:

$$r_n^{(k)} =: \arg\min_{r \geq 0} L_{w,n}[r, \hat{\varepsilon}_n(\cdot; r_n^{(k-1)}), \hat{f}_n(\cdot; r_n^{(k-1)})]. \qquad \square$$

We hope that $\exists \gamma_n := \lim_{k \to \infty} r_n^{(k)}$ and $\gamma_n \approx r_{opt}^{(n,w)}$.

5. Simulation

Data generating process:

$$dX_t = \mu X_t \, dt + \sigma \, dw_t + dz_t$$

where z is a compound Poisson process with Lévy density

$$f(z) = \frac{\lambda}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}.$$

$(\mu, \sigma, \lambda) = (-0.03, 0.3, 15.0)$.

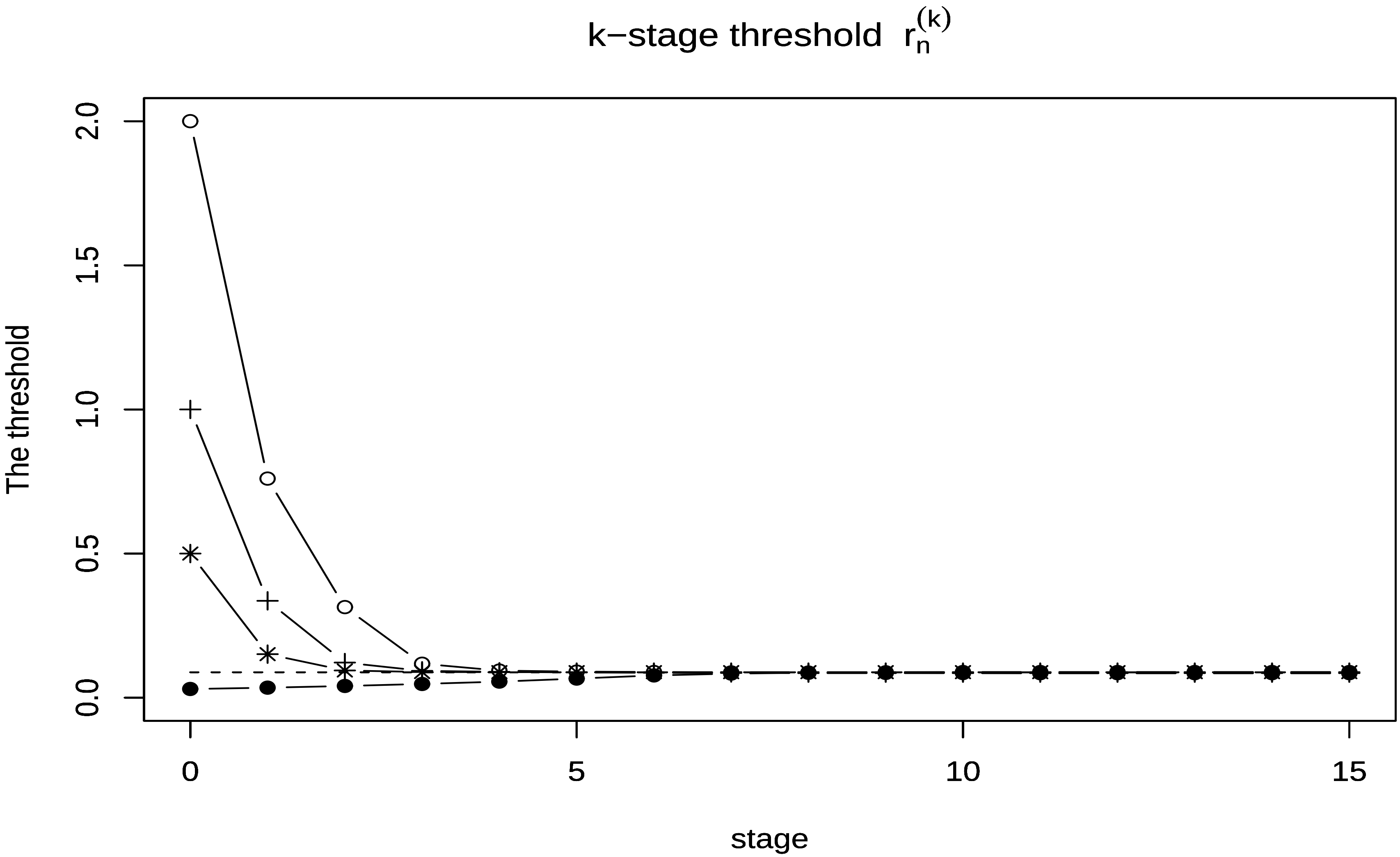
We set $w = 1/2$ and estimated 0-stage's – 7-stage's $r_n^{(k)}, \hat{\mu}_n^{(k)}, \hat{\sigma}_n^{(k)}$ and $\hat{\lambda}_n^{(k)}$ from discrete observations $\boldsymbol{X}^n (h_n = n^{-0.7})$ in an one-path.
Experiments: 500 times.

$n = 1000$	$r_n^{(k)}$	$\hat{\lambda}_n^{(k)}$	$\hat{\mu}_n^{(k)}$	$\hat{\sigma}_n^{(k)}$
0-stage	1.0	4.69	-0.13130	1.72986
s.d.	0.0	1.103	0.322	0.192
1-stage	0.35891	10.52	-0.03974	0.51265
s.d.	0.0515	1.471	0.088	0.085
2-stage	0.09872	13.41	-0.03271	0.30556
s.d.	0.0148	1.788	0.051	0.008
3-stage	0.05819	14.36	-0.03198	0.29740
s.d.	0.0016	1.949	0.051	0.007
4-stage	0.05595	14.60	-0.03180	0.29617
s.d.	0.0018	2.040	0.051	0.007
5-stage	0.05557	14.65	-0.03173	0.29589
s.d.	0.0020	2.050	0.051	0.007
6-stage	0.05548	14.65	-0.03170	0.29580
s.d.	0.0020	2.062	0.051	0.007
7-stage	0.05545	14.67	-0.03172	0.29574
s.d.	0.0020	2.062	0.051	0.007
$r_{opt}^{(n,0.5)}$/true	0.05560	15.0	-0.03	0.3

$n = 3000$	$r_n^{(k)}$	$\hat{\lambda}_n^{(k)}$	$\hat{\mu}_n^{(k)}$	$\hat{\sigma}_n^{(k)}$
0-stage	1.0	4.69	-0.08762	1.73774
s.d.	0.0	0.955	0.223	0.160
1-stage	0.26496	11.74	-0.03145	0.40689
s.d.	0.0225	1.336	0.048	0.040
2-stage	0.05631	14.21	-0.03091	0.30103
s.d.	0.0052	1.588	0.035	0.004
3-stage	0.04092	14.80	-0.03044	0.29908
s.d.	0.0008	1.714	0.035	0.004
4-stage	0.04036	14.90	-0.03023	0.29883
s.d.	0.0008	1.744	0.035	0.004
5-stage	0.04028	14.91	-0.03028	0.29878
s.d.	0.0009	1.758	0.035	0.004
6-stage	0.04027	14.92	-0.03029	0.29878
s.d.	0.0009	1.758	0.035	0.004
7-stage	0.04027	14.92	-0.03029	0.29878
s.d.	0.0009	1.758	0.035	0.004
$r_{opt}^{(n,0.5)}$/true	0.04016	15.0	-0.03	0.3

$n = 10000$	$r_n^{(k)}$	$\hat{\lambda}_n^{(k)}$	$\hat{\mu}_n^{(k)}$	$\hat{\sigma}_n^{(k)}$
0-stage	1.0	4.71	-0.07670	1.73822
s.d.	0.0	0.877	0.179	0.152
1-stage	0.18585	12.66	-0.03042	0.34109
s.d.	0.0153	1.283	0.034	0.016
2-stage	0.03219	14.57	-0.03005	0.29999
s.d.	0.0016	1.442	0.029	0.002
3-stage	0.02777	14.91	-0.02997	0.29942
s.d.	0.0004	1.495	0.029	0.002
4-stage	0.02761	14.95	-0.02995	0.29942
s.d.	0.0004	1.503	0.029	0.002
5-stage	0.02760	14.95	-0.02995	0.29942
s.d.	0.0004	1.504	0.029	0.002
6-stage	0.02760	14.95	-0.02995	0.29942
s.d.	0.0004	1.504	0.029	0.002
7-stage	0.02760	14.95	-0.02995	0.29942
s.d.	0.0004	1.504	0.029	0.002
$r_{opt}^{(n,0.5)}$/true	0.02752	15.0	-0.03	0.3

Stability for the pilot threshold $r_n^{(0)}$



$n = 1000$ and - - -: $r_{opt}^{(n,0.5)} = 0.0556.$

7. Validity of Plug-in method

First, we say that

When $|\Delta_i X^n| > r_n$,

$$\Delta_i X^n \approx \text{Jump size} \quad \Rightarrow \quad \lim_{n\to\infty} r_n = \lim_{n\to\infty} \frac{\sqrt{h_n}}{r_n} = 0.$$

Therefore $r_{opt}^{(n,w)}$ and $\gamma_n := \lim_{k\to\infty} r_n^{(k)}$ should satisfy at least the above necessary conditions.

Under some **regularities** on X, we obtain the follwoing results.

Theorem 1 For each $n \in \mathbf{N}$ and $w \in [1/2, 1]$, there exists a constant γ_n independent of w such that

$$\lim_{k \to \infty} r_n^{(k)} = \gamma_n > 0 \quad a.s.$$

Moreover

$$\lim_{n \to \infty} \left(\gamma_n + \frac{\sqrt{h_n}}{\gamma_n} \right) = 0.$$

Note: Limit γ_n might possibly depend on the pilot threshold $r_n^{(0)}$.

Theorem 2 For any $w \in [1/2, 1]$,

$$\lim_{n\to\infty}\left(r_{opt}^{(n,w)} + \frac{\sqrt{h_n}}{r_{opt}^{(n,w)}}\right) = 0,$$

and, for any $c \in (0, 1/2)$,

$$err_n := \left|r_{opt}^{(n,w)} - \gamma_n\right| < 5h_n^c$$

for any $n \in \mathbf{N}$ *uniformly in the pilot threshold* $r_n^{(0)}$.

Note: Although the limit γ_n might depend on the pilot threshold $r_n^{(0)}$, using γ_n as a threshold is not so unrealistic since err_n is usually very small.

References

[1] Bandi, F. M. and Nguyen, T. H. (2003) On the functional estimation of jump-diffusion models, *J. Econometrics*, **116**, 211–229.

[2] Jiang, J. and Knight, L. (2002) Estimation of continuous-time processes via the empirical characteristic function, *J. Bus. Econom. Statist.*, **20**, no. 2, 198–212.

[3] Mancini, C. (2004) Estimation of the Characteristics of the jumps of a General Poisson-Diffusion model, *Scand. Actuarial J.*, no. 1, 42–52.

[4] Shimizu, Y. (2006) The practical inference for some jump-type stochastic processes from sampled data, Preprint, *Asymptotic Methods in Probability and Statistics*, University of Tokyo, *submitted*.

[5] Shimizu, Y. (2006) Density type estimation of Lévy measure for discretely observed diffusion processes with jumps, *J. Japan Statist. Soc.*, **36**, no. 1, 37-62.

[6] Shimizu, Y. and Yoshida, N. (2006) Estimation of parameters for diffusion processes with jumps from discrete observations, *Statist. Infer. Stochastic Proc.*, **9**, no. 3, 227–277.