Statistical methods for dynamical stochastic models - DYNSTOCH 2016, University of Rennes, June 8th to June 10th

Quasi likelihood analysis and limit order book modeling

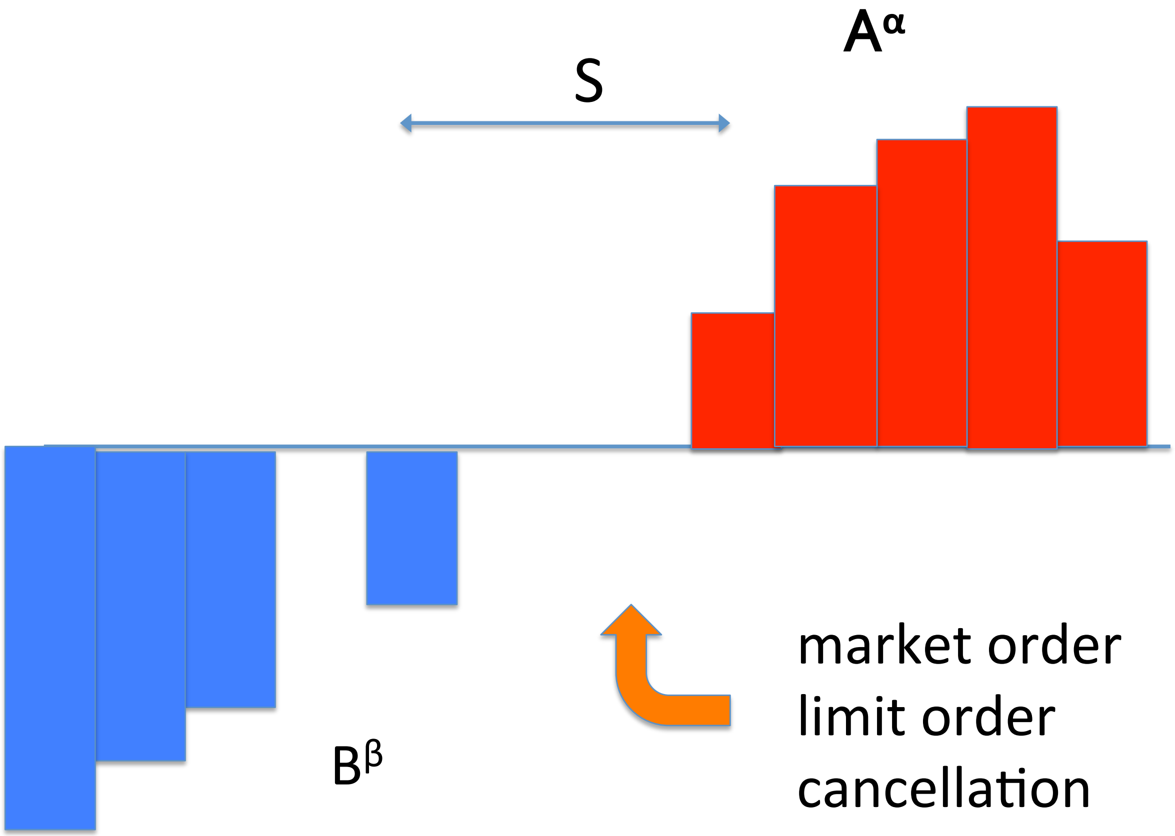
Nakahiro Yoshida

The University of Tokyo
Japan Science and Technology Agency, CREST
Institute of Statistical Mathematics

A Limit Order Book modeling

With Ioane Muni Toke

Limit Order Book modeling:



Limit Order Book modeling

- **LOB is described by the multi-dimensional stochastic process**

$$\mathbb{X} = ((A^{\alpha})_{\alpha=1,\dots,k_A}, (B^{\beta})_{\beta=1,\dots,k_B})$$

where

  - A_t^{α}: **total number of limit orders available at price (tick)** p_A^{α} **on the ask side at time** t
  - B_t^{β}: **total number of limit orders available at price (tick)** p_B^{β} **on the bid side at time** t

- **The state space of** $\mathbb{X}$ **is absolutely or relatively set:**
  - **the price** p_A^{α} **is at the relative** α**-th limit order from the best quote on the same/opposite side,**
    or p_A^{α} **is the absolute price**

- **The random evolution of $\mathbb{X}$ is determined by the processes**
  - **M^A counting number of arrivals of market orders on the ask side,**
  - **M^B of market orders on the bid side,**
  - **L^α of limit orders at level α on the ask side,**
  - **L^β of limit orders at level β on the bid side,**
  - **C^α of cancellation at level α on the ask side, and**
  - **C^β of cancellation at level β on the bid side.**
- **The multivariate counting process N^n consists of these counting processes. Here prices can be recognized as a function of $\mathbb{X}$.**
- **For modeling of C^α and C^β, we may treat $g^n(t, \theta)$ proportional to A^α and B^β, respectively, or more complicated mechanism.**

Counting arrivals of limit orders

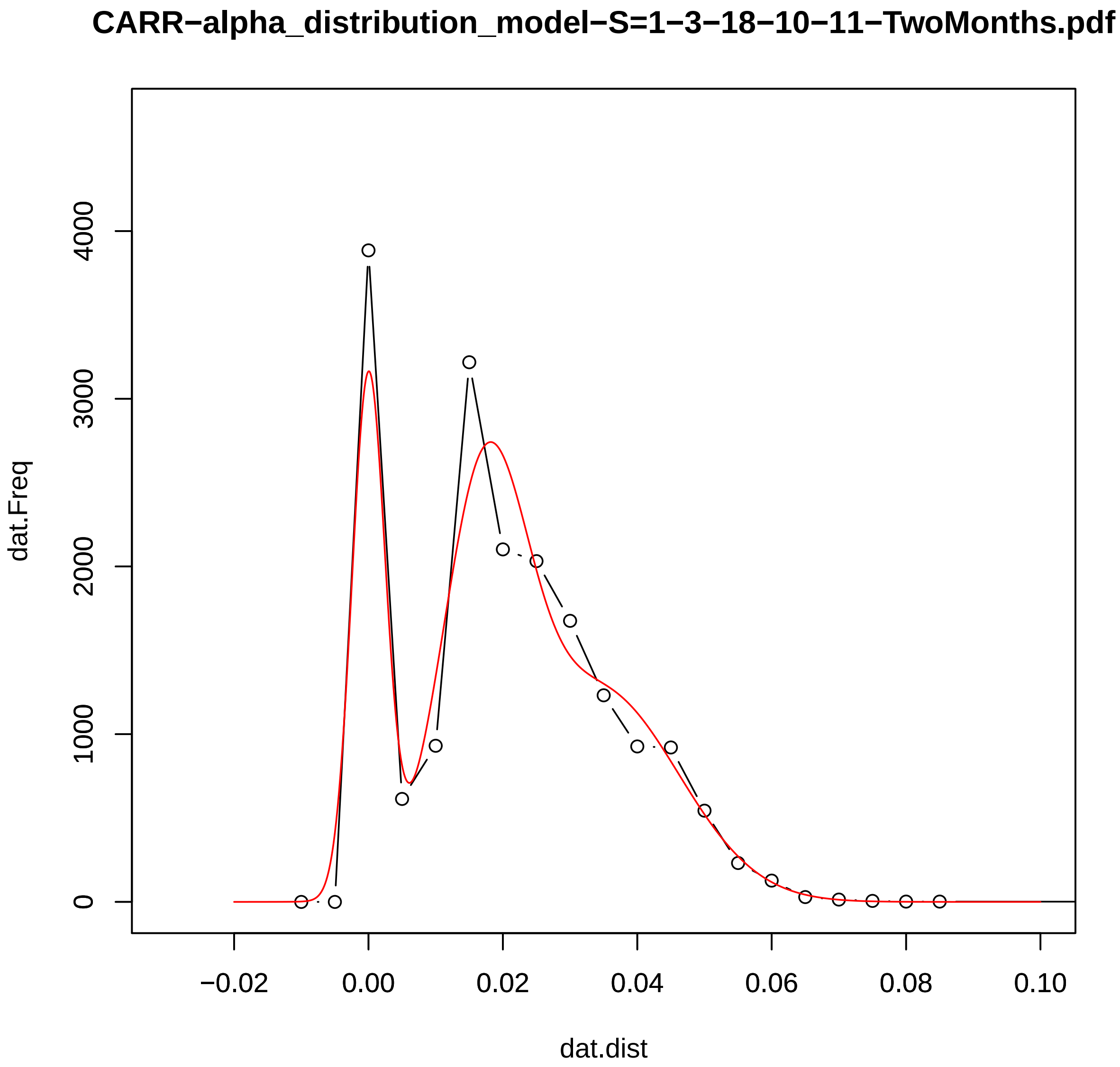


Figure 1: CARR: S-conditional alpha-distribution model (red)

Limit order intensities modeling

- **Modiling limit order intensities λ_α^{LA}**
  - **Discover covariates from the data and give a functional representation of λ_α^{LA}.**
  - **A relatively simple dependency has been found:**

$$\lambda_\alpha^{LA} = \lambda_\alpha^{LA}(S_t)$$

Limit order intensity model (Muni Toke and Y)

- Intensity model (spot form) is proposed as

$$\lambda_\alpha^{LA}(S) = \sum_{i=1}^{3} \Lambda_i(S)\phi(\alpha\delta; \mu_i(S), \mathrm{sd}_i(S)^2) \qquad (\alpha \in \mathbb{R})$$

  – Λ_i are positive functions of S

$$\Lambda_i(S) = \exp(\beta(S))\pi_i(S). \tag{1}$$

  – A model $\beta(s) = \sum_{j=0}^{2} \beta_j s^j$.

$$\mu_i(s) = \sum_{j=0}^{2} \mu_{i,j} s^j \qquad \mathrm{sd}_i(s) = \sum_{j=0}^{2} \sigma_{i,j} s^j$$

$$\pi_i(s) = \frac{\exp\left(\pi_{i,0} + \pi_{i,1}s + \pi_{i,2}s^2\right)}{\sum_{j=1}^{3} \exp\left(\pi_{j,0} + \pi_{j,1}s + \pi_{j,2}s^2\right)}$$

$$(\pi_{3,0} = \pi_{3,1} = \pi_{3,2} = 0)$$

where δ is the tick size, $\mu_{i,j}$, $\sigma_{i,j}$ and $\pi_{i,j}$ are constants depending on the asset and the environment in the sampling period.

Limit order intensity model (by model)

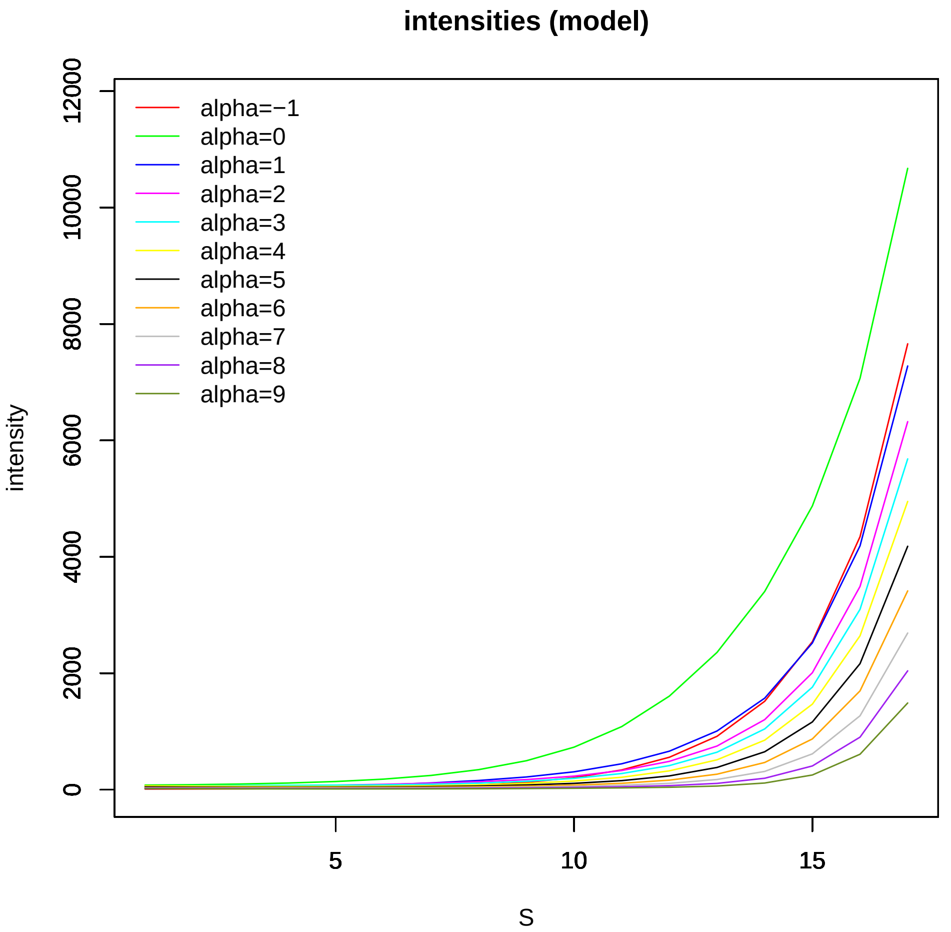


Figure 2: CARR: S and intensities for given $\boldsymbol{\alpha}$ by intensity model

Fit model to LOB arrival numbers data

- **Fit the model to the counting data of the numbers of limit orders for various spreads in a fixed time interval.**
- **Remark. The fitted values are not intensities but the expected numbers of limit orders in the time interval.**

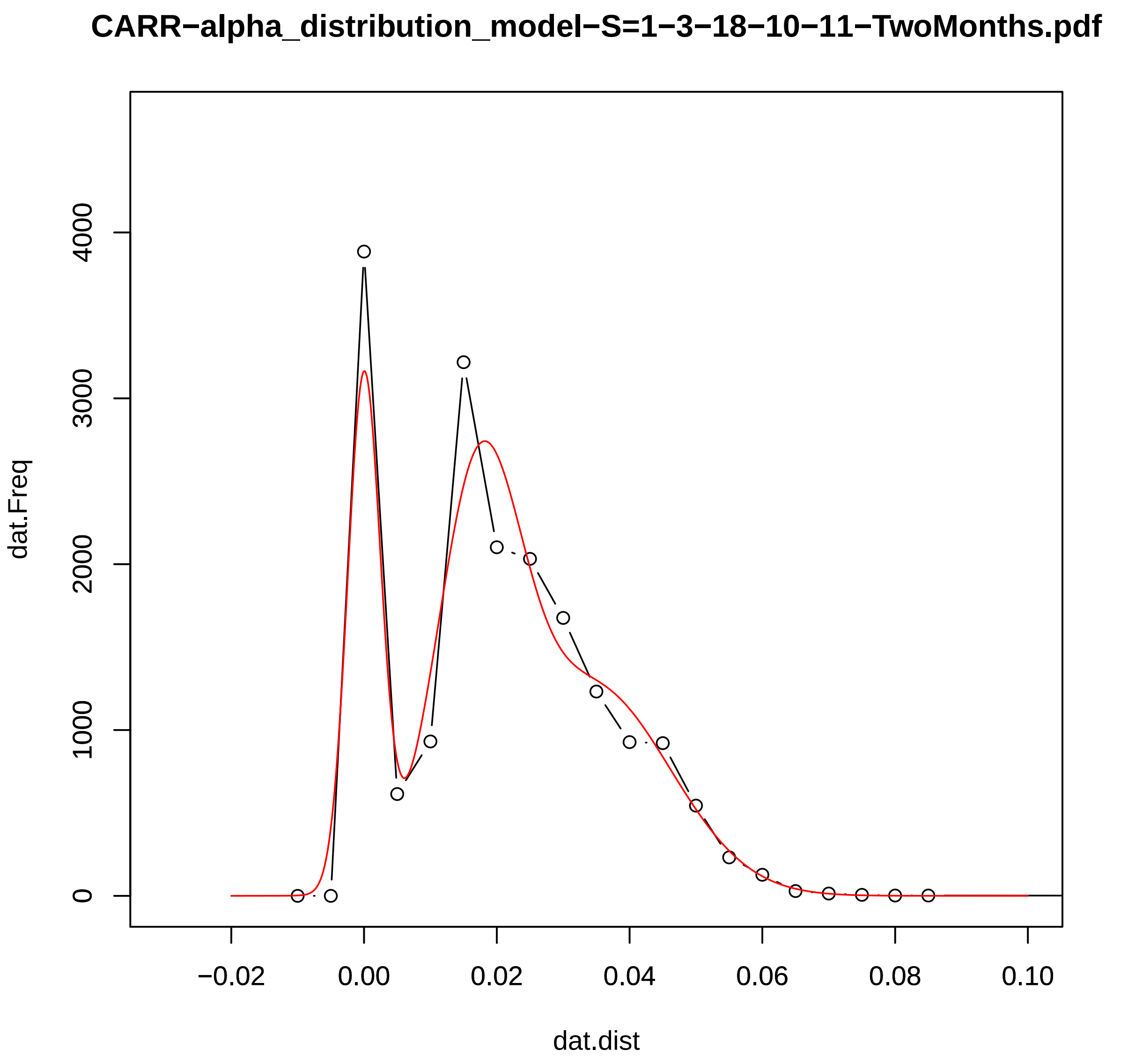


Figure 3: CARR: S-conditional alpha-distribution model (red)

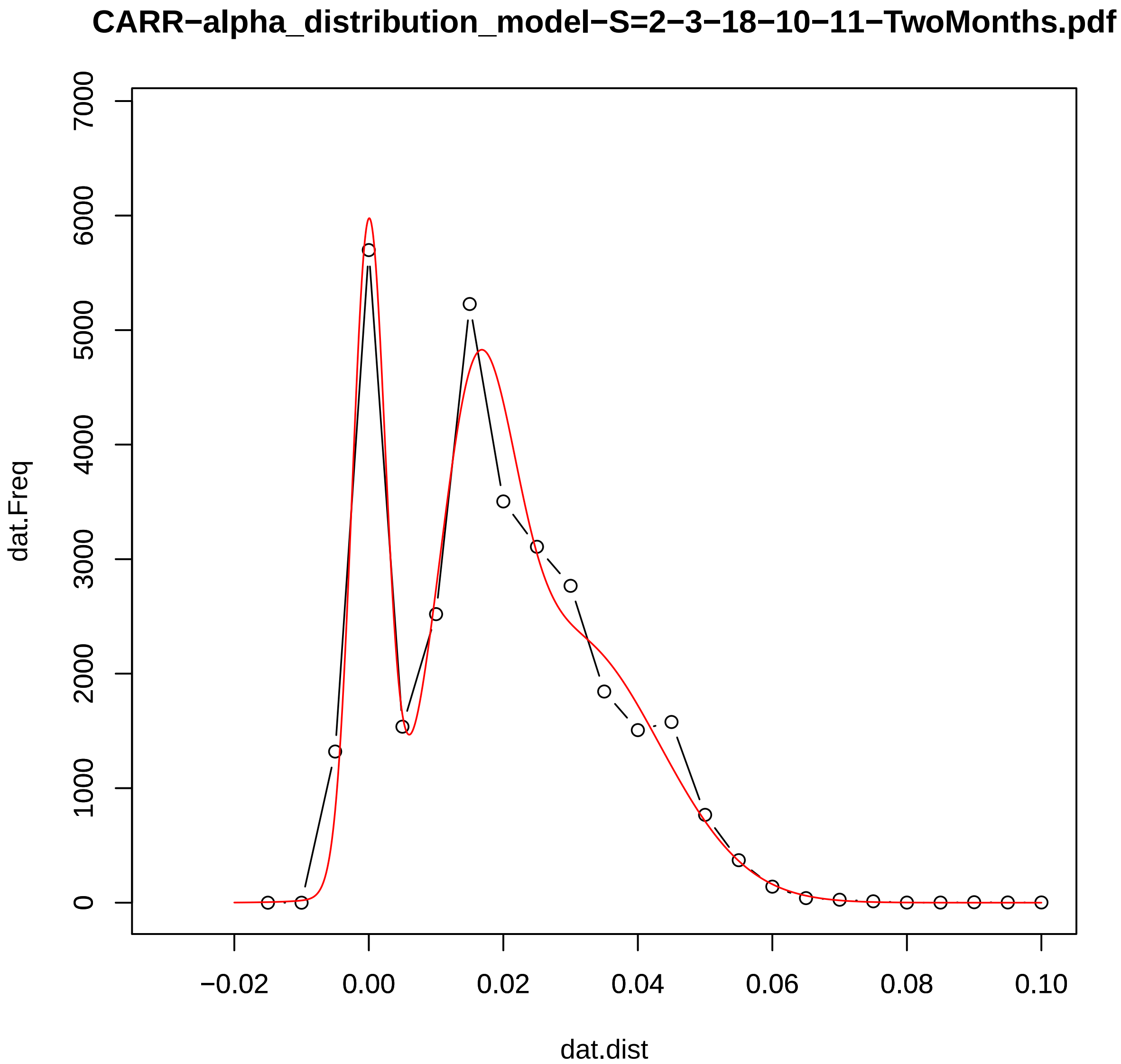


Figure 4: CARR: S-conditional alpha-distribution model (red)

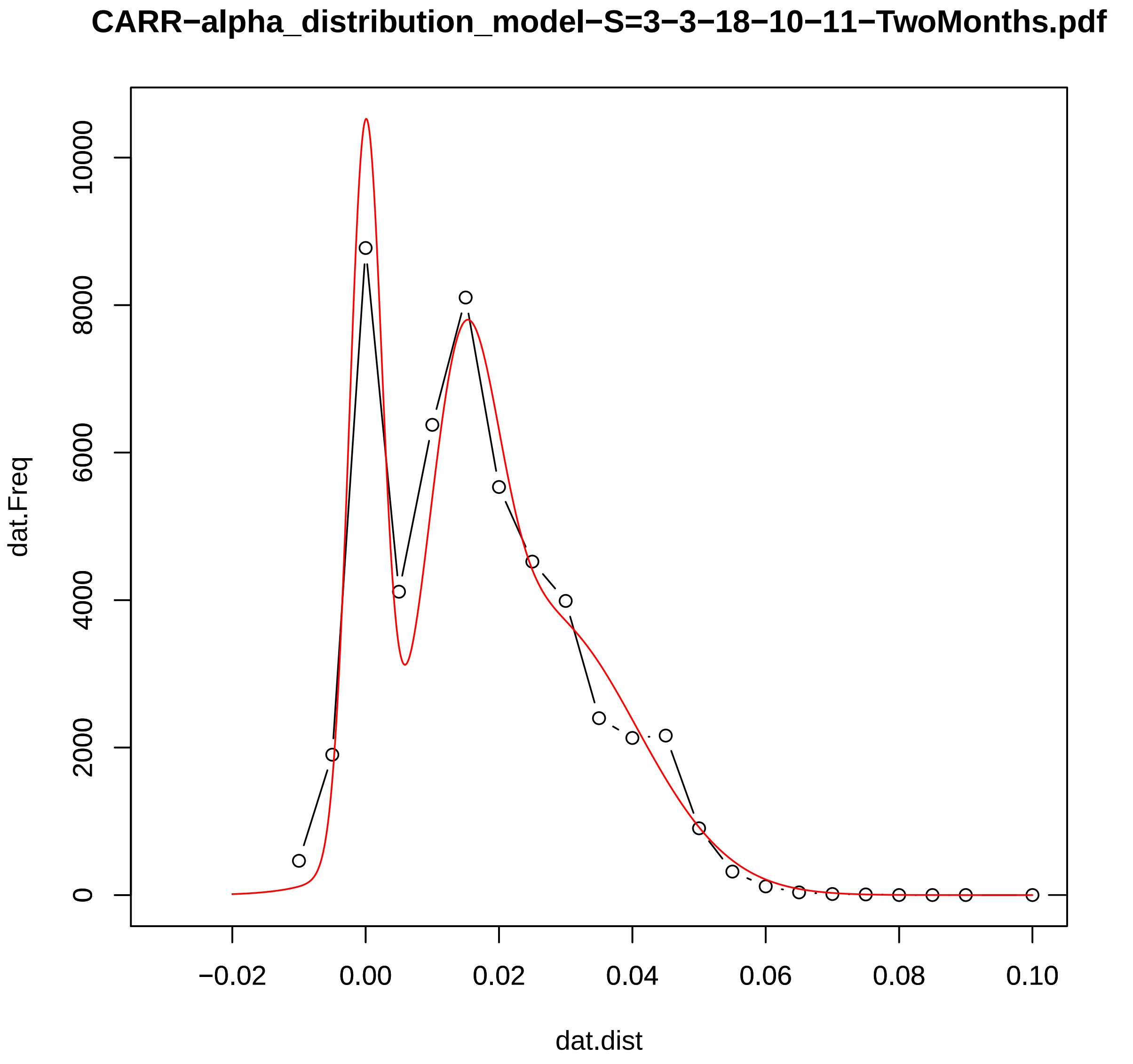


Figure 5: CARR: S-conditional alpha-distribution model (red)

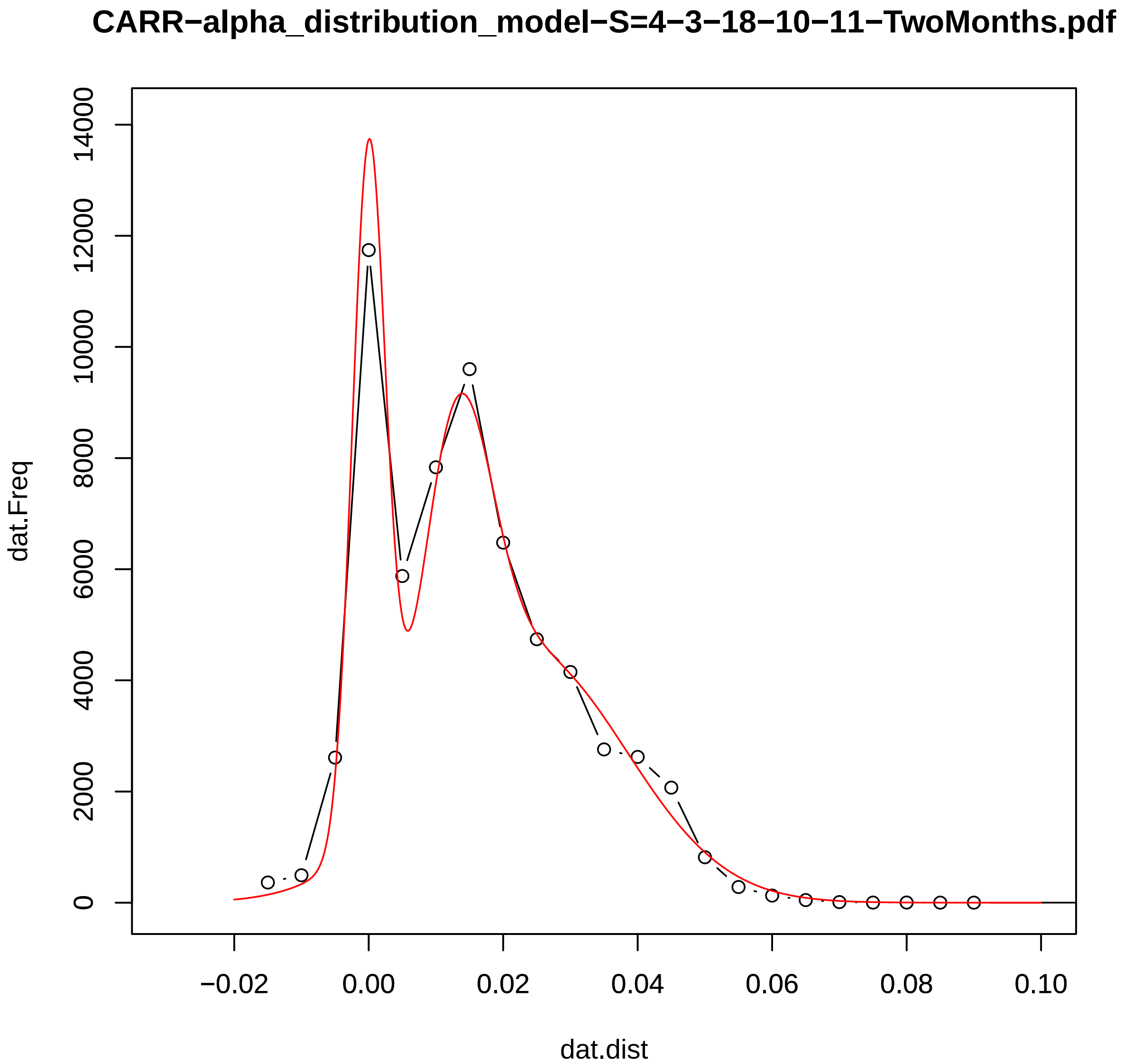


Figure 6: CARR: S-conditional alpha-distribution model (red)

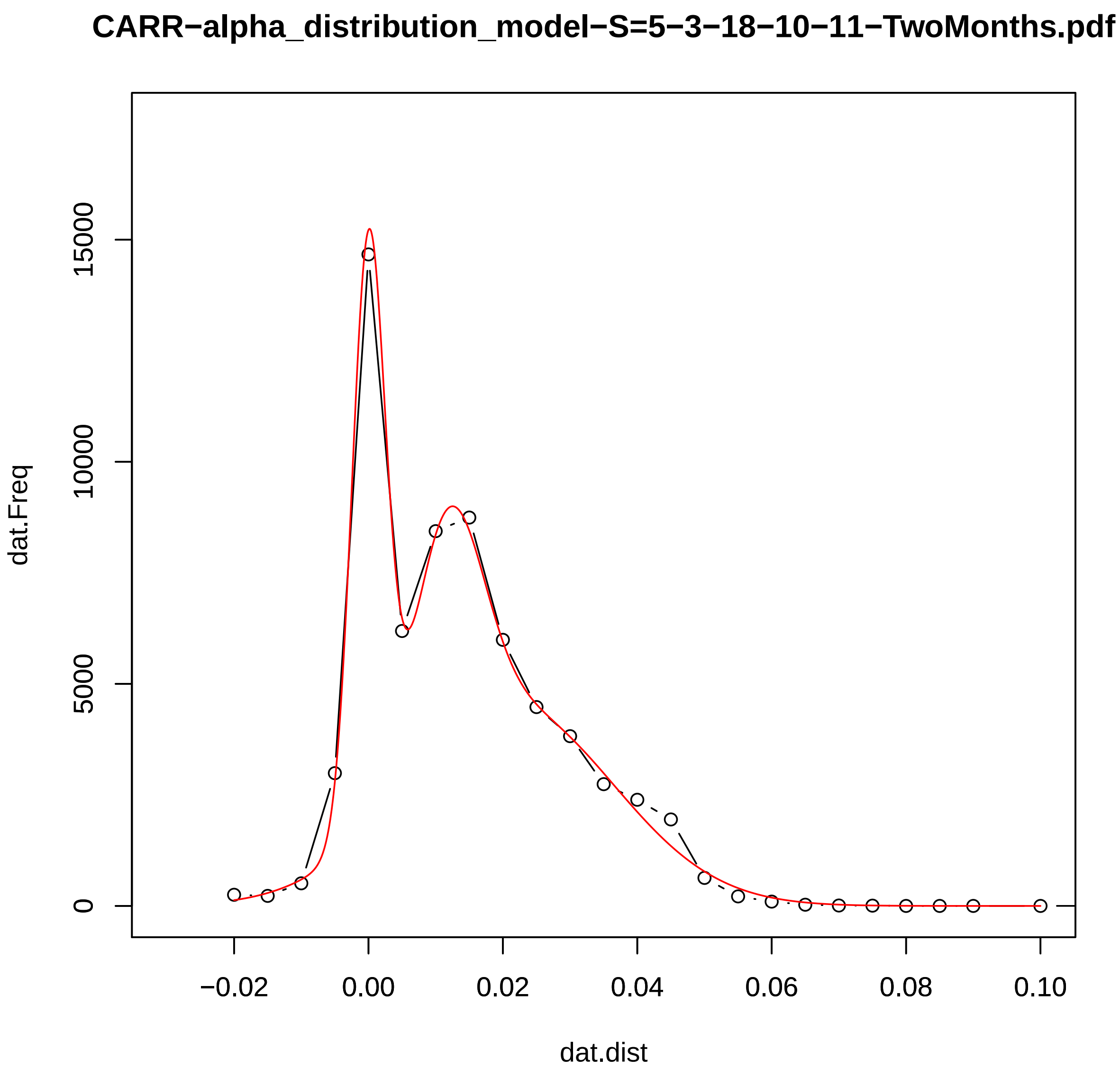


Figure 7: CARR: S-conditional alpha-distribution model (red)

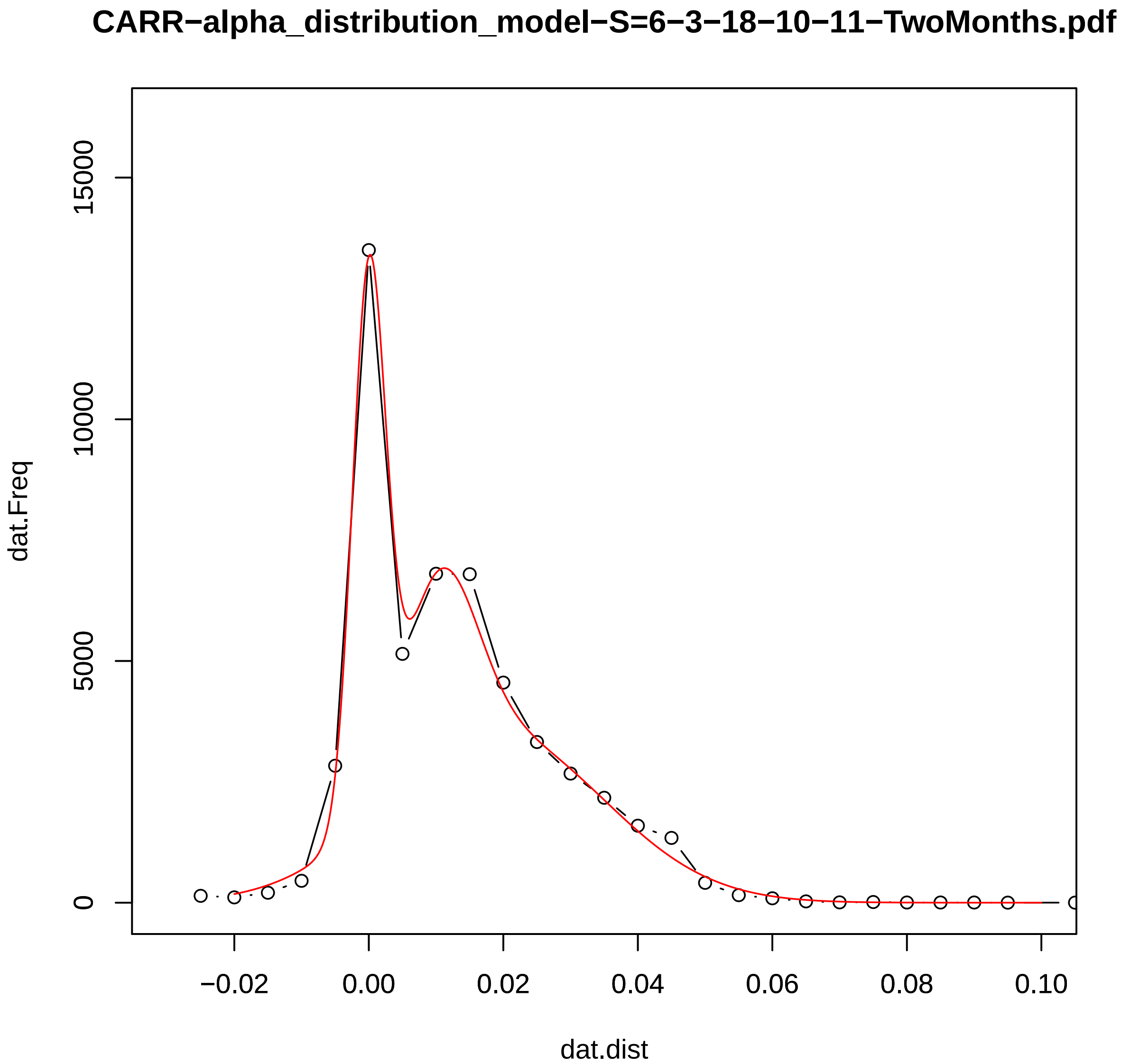


Figure 8: CARR: S-conditional alpha-distribution model (red)

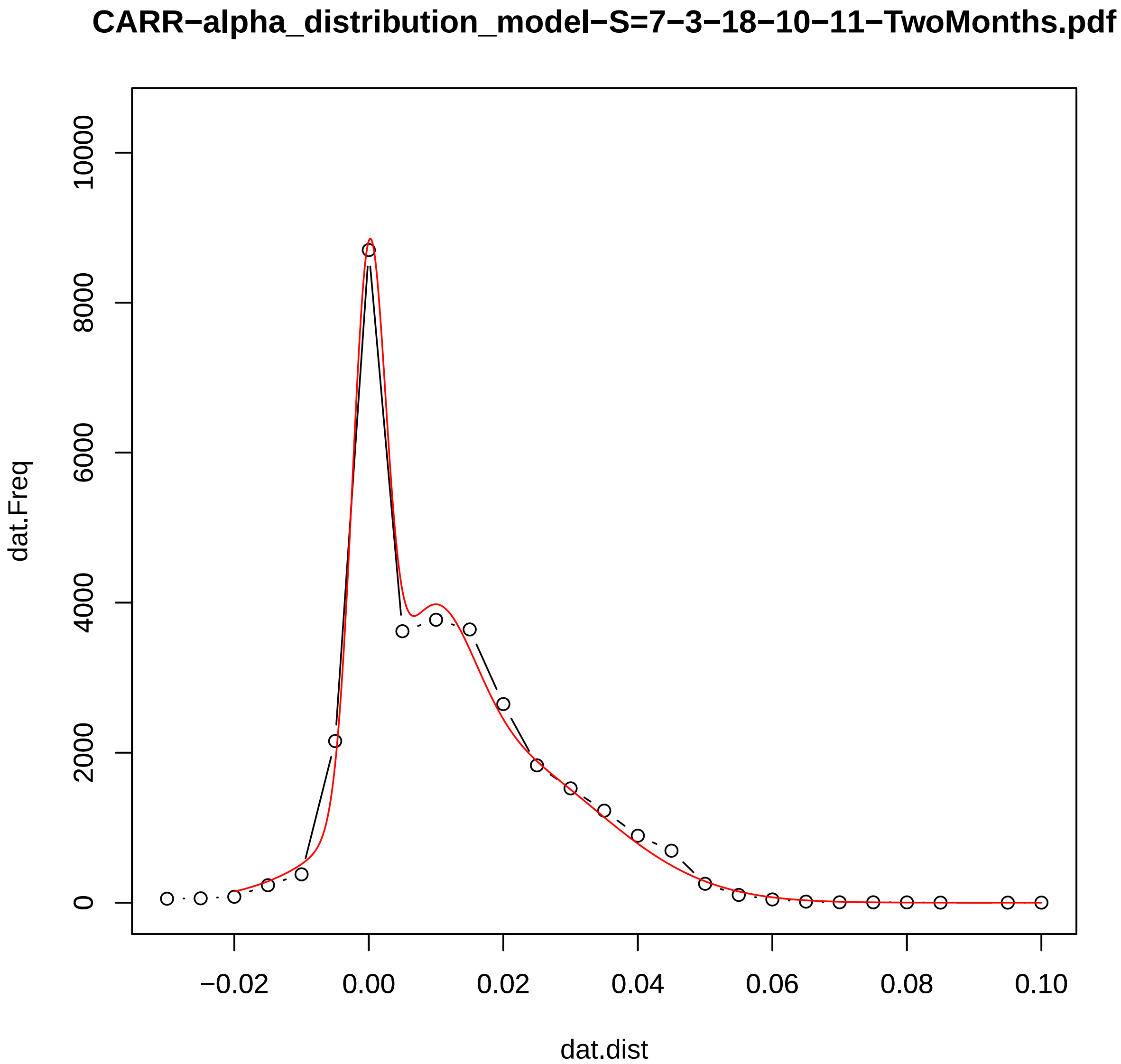


Figure 9: CARR: S-conditional alpha-distribution model (red)

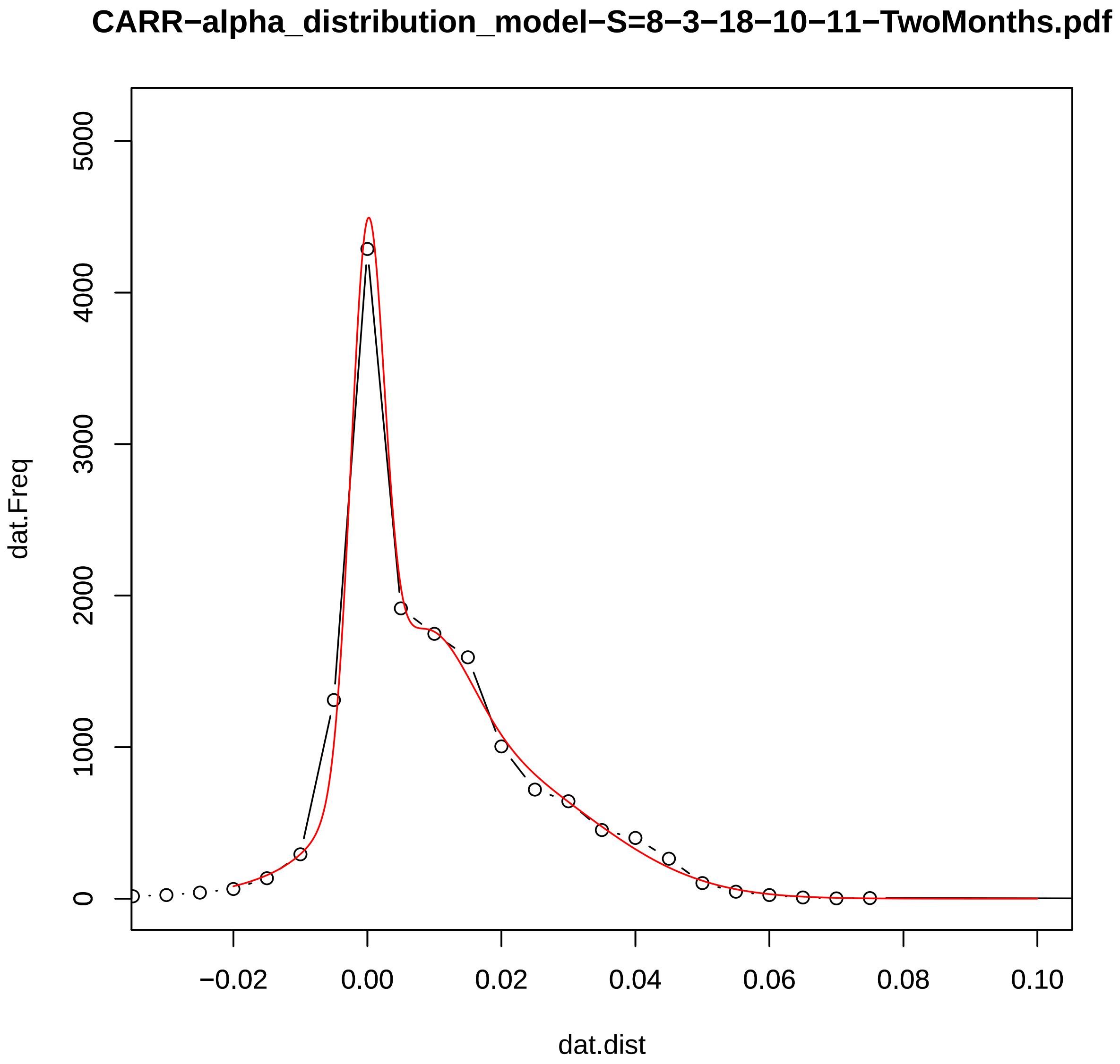


Figure 10: CARR: S-conditional alpha-distribution model (red)

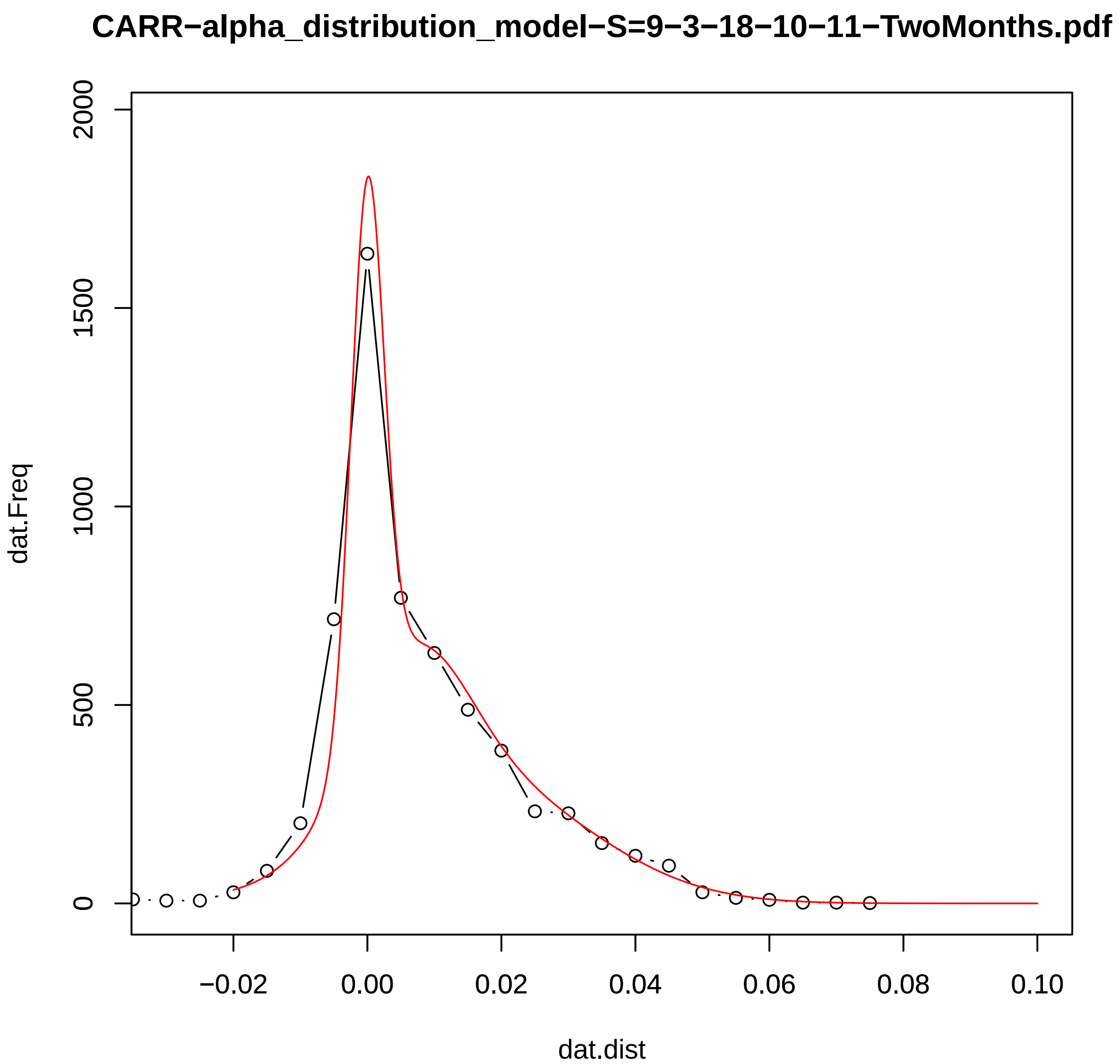


Figure 11: CARR: S-conditional alpha-distribution model (red)

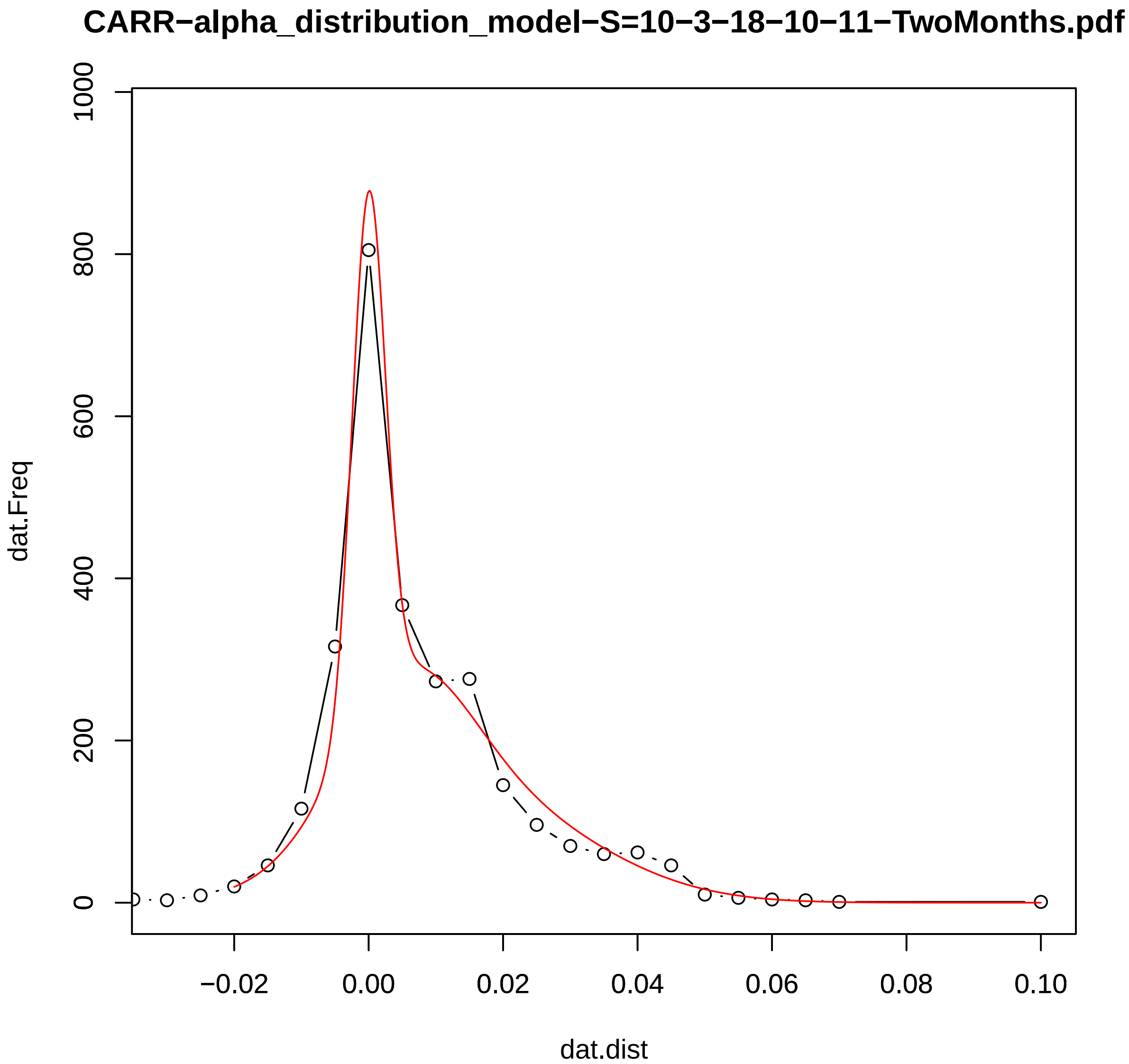


Figure 12: CARR: S-conditional alpha-distribution model (red)

Limit order book modeling

- **This analysis shows possibility of regression model with covariate processes.**

Ultra high frequency data and modeling by point processes

High frequency financial data

hpq_dis_12_2010 - メモ帳

ファイル(F) 編集(E) 書式(O) 表示(V) ヘルプ(H)

HPQ	12/01/2010	N	15:59:54	42.5800	700	F	0	0
HPQ	12/01/2010	N	15:59:54	42.5800	100	@	0	0
HPQ	12/01/2010	N	15:59:55	42.5800	100	@	0	0
HPQ	12/01/2010	N	15:59:55	42.5700	200	@	0	0
HPQ	12/01/2010	N	15:59:55	42.5800	394	@	0	0
HPQ	12/01/2010	N	15:59:55	42.5700	100	@	0	0
HPQ	12/01/2010	N	15:59:55	42.5800	100	@	0	0
HPQ	12/01/2010	N	15:59:57	42.5800	180	@	0	0
HPQ	12/01/2010	N	15:59:59	42.5800	6300	@	0	0
HPQ	12/01/2010	N	15:59:59	42.5800	100	@	0	0
HPQ	12/01/2010	N	15:59:59	42.5700	1000	@	0	0
HPQ	12/01/2010	N	15:59:59	42.5700	6200	@	0	0
HPQ	12/01/2010	N	15:59:59	42.5700	100	@	0	0
HPQ	12/01/2010	N	16:00:07	42.5700	1044254	6	0	0
DIS	12/01/2010	N	9:30:12	37.2300	182284	O	0	0
DIS	12/01/2010	N	9:30:12	37.2100	200	@	0	0
DIS	12/01/2010	N	9:30:12	37.2100	100	@	0	0
DIS	12/01/2010	N	9:30:12	37.2100	100	@	0	0
DIS	12/01/2010	N	9:30:12	37.2100	202	@	0	0
DIS	12/01/2010	N	9:30:12	37.2200	400	F	0	0
DIS	12/01/2010	N	9:30:13	37.1900	300	@	0	0
DIS	12/01/2010	N	9:30:15	37.2300	288	@	0	0
DIS	12/01/2010	N	9:30:15	37.2400	200	F	0	0
DIS	12/01/2010	N	9:30:15	37.2500	712	F	0	0
DIS	12/01/2010	N	9:30:15	37.2500	108	@	0	0
DIS	12/01/2010	N	9:30:15	37.2000	100	@	0	0
DIS	12/01/2010	N	9:30:16	37.2500	1584	@	0	0

Phenomena we want to model

- Epps effect (1979)
  A natural correlation estimator has a bias in high frequent observations
  - non-synchronicity of the observations
  - microstructure Observations are $X_{t_j} + \epsilon_j$?
    No BM in ultra high frequency sampling
- lead-lag
- relativity of prices — In Limit order Book,
  "Price" is a functional of the state of LOB.
- Dependency on covariates
- To incorporate these effects, we will consider
  a point process regression model.

Modeling high frequency data by point processes

- **Multivariate point process**
  - **Hewlett (2006)**
    the clustered arrivals of buy and sell trades using Hawkes processes
  - **Large (2007)**
    Extension by using a finer description of orders
  - **Bowsher (2007)**
    Generalized Hawkes model
  - **E. Bacry et al. (2013)**
    Price as "upward − downward counting processes"
  - **Chen and Hall (2013)**
    the intraday trading times of a common stock traded on the Australian Stock Exchange, the ANZ stock.

Modeling high frequency data by point processes

- **Limit order book**
  - **R. Cont, Stoikov and Talreja (2010)**
  - **Abergel and Jedidi (2013)**
  - **Smith, Farmer, Gillemot and Krishnamurthy (2003)**
  - **Muni Toke and Pomponio (2011)**

Point process regression model

Ogihara and Yoshida, arXiv 2015

Point process regression model

- The **d**-dimensional point process $N^n = (N^{n,\alpha})_{\alpha \in \mathcal{I}}$ on $I = [T_0, T_1]$, $\mathcal{I} = \{1, ..., \mathbf{d}\}$, is assumed to have

an intensity process $n\lambda^n(t, \theta)$ defined by

$$\lambda^n(t, \theta) = g^n(t, \theta) + \int_{\hat{T}_0}^{t-} K^n(t, s, \theta) dX_s^n,$$

where θ is a parameter and $\hat{T}_0 < T_0 < T_1$.

- Examples.

$$\lambda^n(t, \theta) = \lambda^\infty(t, \theta) = g(V_t, \theta)$$

$$\lambda^n(t, \theta) = \lambda^\infty(t, \theta) = g(t, \gamma) + \int_0^t e^{-b(t-s)} A V_s ds$$

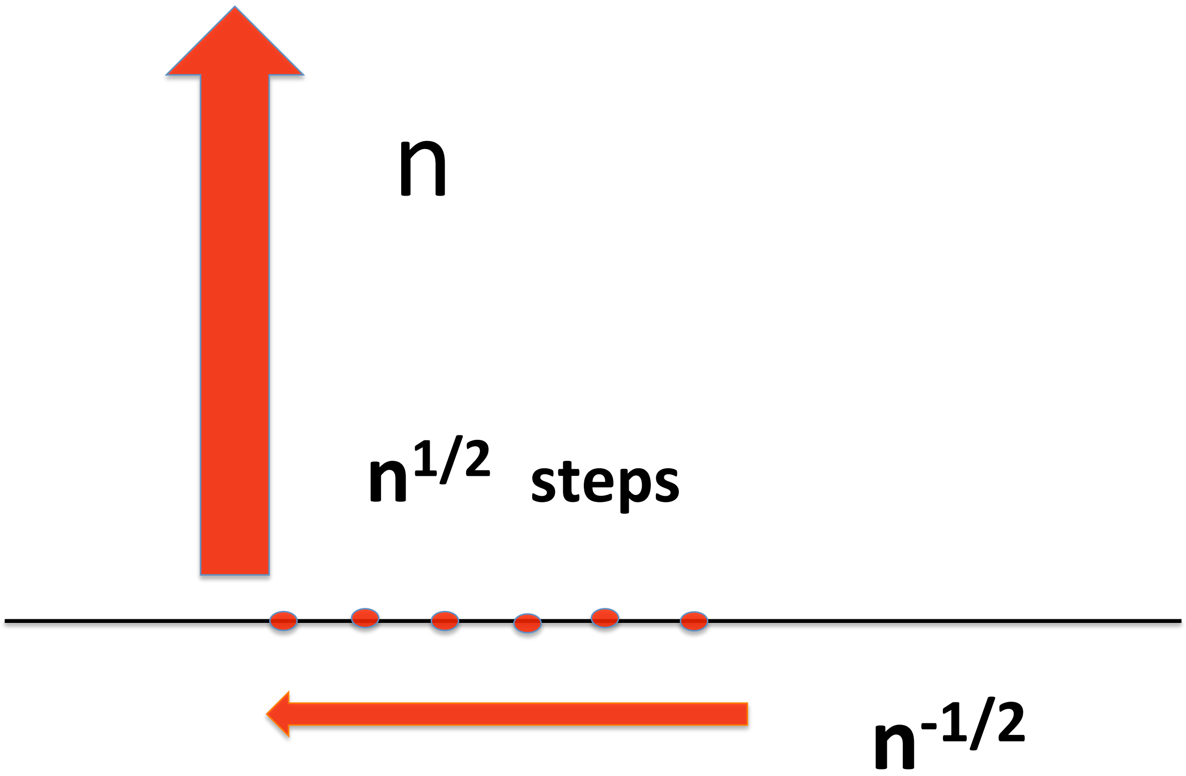
for a random covariate process V_t.

- More precisely, we will work on
  - a stochastic basis $\mathcal{B} = (\Omega, \mathcal{F}, \mathrm{F}, P)$,
  - $\mathrm{F} = (\mathcal{F}_t)_{t \in \hat{I}}$ being a filtration on $(\Omega, \mathcal{F})$, where $\hat{I} = [\hat{T}_0, T_1] \supset I$ and $n \in \mathbb{N}$.
  - For each $n \in \mathbb{N}$ and $\theta \in \Theta$, $(g^n(t, \theta))_{t \in I}$ is a d-dimensional predictable process,
  - $(K^n(t, s, \theta))_{s \in [\hat{T}_0, t)}$ is a $\mathsf{d} \times \mathsf{d}_0$ matrix-valued optional process for $t \in I$, $\mathcal{I}_0 = \{1, ..., \mathsf{d}_0\}$, and
  - $(X^n_t)_{t \in \hat{I}}$ is a d_0-dimensional F-adapted right-continuous increasing process on $\mathcal{B}$.
- The multivariate point process N^n is compensated by the process $(\int_{T_0}^t n\lambda^n(s, \theta) ds)_{t \in I}$ when θ is the true value of the unknown parameter.
- No common jumps of different elements of N^n

Our model

- **regression of the intensities to covariate processes and their history**
- **finite time horizon and the intensities of point processes tends to ∞ —- non-ergodic statistics**

Locally Poissonian and Globally non-ergodic model



Local model such as $\lambda_\alpha^{LA}(S_t)$ becomes a fibre.

Quasi Likelihood Analysis (QLA)

- **Ibragimov-Hasminskii and Kutoyants' program**
  + polynomial type large deviation inequality
  = Quasi likelihood analysis:
  - **ergodic / non-ergodic**
  - **limit theorems for QMLE and QBE**
  - **convergence of moments**
  - **Y. AISM 2011**

Recall Point Process Regression Model

$\boldsymbol{N^n = (N^{n,\alpha})_{\alpha \in \mathcal{I}}}$ **has an intensity process** $\boldsymbol{n\lambda^n(t,\theta)}$ **defined by**

$$\boldsymbol{\lambda^n(t,\theta) = g^n(t,\theta) + \int_{\hat{T}_0}^{t-} K^n(t,s,\theta)dX_s^n,}$$

Quasi likelihood

- **We shall consider estimation for the unknown parameter θ.**
- **Observations**

$$(N_t^{n,\alpha})_{t\in I,\alpha\in\mathcal{I}}, \quad (X_t^{n,\beta})_{t\in\hat{I},\beta\in\mathcal{I}_0},$$
$$(g^{n,\alpha}(t,\theta))_{t\in I,\alpha\in\mathcal{I},\theta\in\Theta},$$
$$(K_\beta^{n,\alpha}(t,s,\theta))_{t\in I,s\in[\hat{T}_0,t),\alpha\in\mathcal{I},\beta\in\mathcal{I}_0,\theta\in\Theta}.$$

This is the case, for example, when $g^{n,\alpha}(t,\theta)$ is a function of θ and some observable covariate process:

$$g^{n,\alpha}(t,\theta) = g^{\alpha}(t,V_t,\theta) \quad \text{with observable } V_t$$

- **The quasi log likelihood is given by**

$$\boldsymbol{l_n(\theta) = \sum_{\alpha \in \mathcal{I}} \left(\int_{T_0}^{T_1} \log[n\lambda^{n,\alpha}(t,\theta)] dN_t^{n,\alpha} - \int_{T_0}^{T_1} [n\lambda^{n,\alpha}(t,\theta) - 1] dt \right)}$$

for observed point process N^n. Obviously, "-1" in the second integral can be eliminated for maximization. The factor "n" in the first integral is also unnecessary. Thus we can use

$$\boldsymbol{\ell_n(\theta) = \sum_{\alpha \in \mathcal{I}} \left(\int_{T_0}^{T_1} \log \lambda^{n,\alpha}(t,\theta) dN_t^{n,\alpha} - \int_{T_0}^{T_1} n\lambda^{n,\alpha}(t,\theta) dt \right)} \tag{2}$$

instead of $l_n(\theta)$.

Statistical random field $\mathbb{Z}_n$

We shall work with the statistical random field

$$\mathbb{H}_n(\theta) = \ell_n(\theta)$$

on Θ and apply the frame of the quasi likelihood analysis.The random fields $\mathbb{Z}_n$ is defined on $\mathbb{U}_n = \{u \in \mathbb{R}^{\mathsf{p}};\ \theta_u \in \Theta\}$, $\theta_u = \theta^* + n^{-1/2}u$, by

$$\begin{aligned} \mathbb{Z}_n(u) &= \exp\big(\mathbb{H}_n(\theta_u) - \mathbb{H}_n(\theta^*)\big) \\ &= \exp\Bigg(\sum_{\alpha=1}^{\mathsf{d}} \int_{T_0}^{T_1} \log \frac{\lambda^{n,\alpha}(t,\theta_u)}{\lambda^{n,\alpha}(t,\theta^*)}\, dN_t^{n,\alpha} \\ &\qquad - \sum_{\alpha=1}^{\mathsf{d}} \int_{T_0}^{T_1} n\big[\lambda^{n,\alpha}(t,\theta_u) - \lambda^{n,\alpha}(t,\theta^*)\big]dt\Bigg). \end{aligned}$$

Assumptions

- **Assume that the boundary of Θ is good and that the function $\Theta \ni \theta \mapsto \lambda^n(t, \theta)$ has continuous extension to $\bar{\Theta}$ when the QMLE is discusses.**
- **Let ε be a positive number less than 1/2.**
- $\bar{\mathbb{N}} := \mathbb{N} \cup \{\infty\}$

Assumptions

[B1]$_{\bar{j}}$ For each $n \in \bar{\mathbb{N}}$, $K^n(t,s,\theta)$ is an $\mathbb{R}_+^{\mathsf{d}} \otimes \mathbb{R}_+^{\mathsf{d}_0}$-valued $\mathcal{F} \times \mathbb{B}(J) \times \mathbb{B}(\Theta)$-measurable function satisfying the following conditions.

(i) For each $(n,t,\theta) \in \mathbb{N} \times I \times \Theta$, the process $[\hat{T}_0, t) \ni s \mapsto K^n(t,s,\theta)$ is $(\mathcal{F}_s)_{s\in[\hat{T}_0,t)}$-optional.

(ii) For each $(n,t,s) \in \bar{\mathbb{N}} \times J$, the mapping $\Theta \ni \theta \mapsto K^n(t,s,\theta)$ is $\bar{j}$ times differentialble a.s., $\sup_{(s,\theta)\in[\hat{T}_0,t)\times\Theta} |\partial_\theta^j K^n(t,s,\theta)| < \infty$ a.s. for $t \in I$, and

$$\sum_{j=0}^{\bar{j}} \sup_{(n,s,t)\in\bar{\mathbb{N}}\times J} \sup_{\theta\in\Theta} \|\partial_\theta^i K^n(t,s,\theta)\|_p < \infty$$

for every $p > 1$.

(iii) For each $(n,t,\theta) \in \mathbb{N} \times I \times \Theta$, the mappings $[\hat{T}_0, t) \ni s \mapsto \partial_\theta^i K^n(t,s,\theta)$ $(i = 0,1)$ are differen-tialble a.s., $\sup_{(s,\theta)\in[\hat{T}_0,t)\times\Theta} |\partial_s \partial_\theta^i K^n(t,s,\theta)| < \infty$ a.s. for $t \in I$, and

$$\sup_{(n,t,\theta)\in\mathbb{N}\times I\times\Theta} \sum_{i=0}^{1} \int_{\hat{T}_0}^{t} \|\partial_s \partial_\theta^i K^n(t,s,\theta)\|_p ds < \infty$$

for every $p > 1$.

(iv) For every $p > 1$,

$$n^\varepsilon \sum_{j=0}^{1} \sup_{(t,s)\in J,\theta\in\Theta} \left\| \partial_\theta^j K^n(t,s,\theta) - \partial_\theta^j K^\infty(t,s,\theta) \right\|_p \to 0$$

as $n \to \infty$.

[B2]$_{\bar{j}}$ For each $(\alpha, n) \in \mathcal{I} \times \bar{\mathbb{N}}$, $g^{n,\alpha}(t,\theta)$ is an nonnegative $\mathcal{F} \times \mathbb{B}(I) \times \mathbb{B}(\Theta)$-measurable function for which the following conditions are fulfilled.

(i) For each $(n, \alpha, \theta) \in \mathbb{N} \times \mathcal{I} \times \Theta$, the process $(g^{n,\alpha}(t,\theta))_{t \in I}$ is predictable.

(ii) For each $(n,t) \in \bar{\mathbb{N}} \times I$, the mapping $\Theta \ni \theta \mapsto g^n(t,\theta)$ is $\bar{j}$ times differentiable a.s. and

$$\sum_{j=0}^{\bar{j}} \sup_{(n,t) \in \bar{\mathbb{N}} \times I} \sup_{\theta \in \Theta} \left\| (\partial_\theta)^j g^n(t,\theta) \right\|_p < \infty$$

for every $p > 1$.

(iii) For every $p > 1$,

$$n^\varepsilon \sum_{j=0}^{1} \sup_{t \in I} \sup_{\theta \in \Theta} \left\| \partial_\theta^j g^n(t,\theta) - \partial_\theta^j g^\infty(t,\theta) \right\|_p \to 0$$

as $n \to \infty$.

[B3] For each $n \in \mathbb{N}$ and $\beta \in \mathcal{I}_0$, $(X_t^{n,\beta})_{t\in\hat{I}}$ is a non-decreasing $(\mathcal{F}_t)_{t\in\hat{I}}$-adapted process, and for each $\beta \in \mathcal{I}_0$, there exists a non-decreasing process $(X_t^{\infty,\beta})_{t\in I}$ such that

$$\sup_{(n,t)\in\mathbb{N}\times\hat{I}} \left\|X_t^{n,\beta}\right\|_p < \infty \text{ and}$$
$$n^{\varepsilon} \sup_{t\in\hat{I}} \left\|X_t^{n,\beta} - X_t^{\infty,\beta}\right\|_p \to 0$$

as $n \to \infty$, for every $p > 1$. ($\hat{I} = [\hat{T}_0, T_1]$.)

[B4] For each $(\omega, n, \alpha, t, \theta) \in \Omega \times \mathbb{N} \times \mathcal{I} \times I \times \Theta$, $\lambda^{n,\alpha}(t,\theta) = 0$ if and only if $\lambda^{n,\alpha}(t,\theta) = 0$, and

$$\sup_{(n,t,\theta)\in I\times\Theta} \left\|\lambda^{n,\alpha}(t,\theta)^{-1} 1_{\{\lambda^{n,\alpha}(t,\theta)\neq 0\}}\right\|_p < \infty$$

for every $p > 1$ and $\alpha \in \mathcal{I}$.

Information matrix and Limit intensity process

- Let

$$\Gamma = \sum_{\alpha \in \mathcal{I}} \int_{T_0}^{T_1} (\partial_\theta \lambda^{\infty,\alpha})^{\otimes 2} (\lambda^{\infty,\alpha})^{-1} (t, \theta^*) dt,$$

  where

$$\lambda^{\infty,\alpha}(t,\theta) = g^{\infty,\alpha}(t,\theta) + \int_{\hat{T}_0}^{t-} K_\beta^{\infty,\alpha}(t,s,\theta) dX_s^{\infty,\beta} \tag{3}$$

  for $t \in I$.

- $\lambda^{\infty,\alpha}(t,\theta)$ is possibly random.

Key index

$$\mathbb{Y}(\theta) := -\sum_{\alpha=1}^{\mathsf{d}} \int_{T_0}^{T_1} \Big[\lambda^{\infty,\alpha}(t,\theta) - \lambda^{\infty,\alpha}(t,\theta^*) - \log \frac{\lambda^{\infty,\alpha}(t,\theta)}{\lambda^{\infty,\alpha}(t,\theta^*)} \lambda^{\infty,\alpha}(t,\theta^*)\Big] dt \tag{4}$$

$$\chi_0 := \inf_{\theta\in\Theta\setminus\{\theta^*\}} \frac{-\mathbb{Y}(\theta)}{|\theta-\theta^*|^2}.$$

The nondegeneracy of the key index χ_0 will play an essential role in our argument.

[B5] For every $L > 0$, there exists a constant C_L such that

$$P[\chi_0 < r^{-1}] \leq \frac{C_L}{r^L} \quad (\forall r > 0).$$

Polynomial type large deviation inequality

Theorem 1. (Polynomial type large deviation inequality) Suppose that Conditions $[B1]_4$, $[B2]_4$, $[B3]$, $[B4]$ and $[B5]$ are fulfilled. Then, for every $L > 0$, there exists a constant C_L such that

$$P\left[\sup_{u\in\mathbb{V}_n(r)} \mathbb{Z}_n(u) \geq e^{-r}\right] \leq \frac{C_L}{r^L}$$

for all $r > 0$ and all $n \in \mathbb{N}$, where $\mathbb{V}_n(r) = \{u \in \mathbb{U}_n;\ |u| \geq r\}$.

Quasi likelihood analysis

Denote by $C_{\uparrow}(\mathbb{R}^{\mathsf{p}})$ the set of continuous functions $f : \mathbb{R}^{\mathsf{p}} \longrightarrow \mathbb{R}$ at most polynomial growth.

Theorem 2. Suppose that Conditions $[B1]_4$, $[B2]_4$, $[B3]$, $[B4]$ and $[B5]$ are fulfilled. Then

(a) $\sqrt{n}(\hat{\theta}_n - \theta^*) \longrightarrow^{d_s} \Gamma^{-1/2}\zeta$ **as** $n \to \infty$.

(b) $E\big[f\big(\sqrt{n}(\hat{\theta}_n - \theta^*)\big)\big] \longrightarrow \mathbb{E}\big[f\big(\Gamma^{-1/2}\zeta\big)\big]$ **as** $n \to \infty$ **for all** $f \in C_{\uparrow}(\mathbb{R}^{\mathsf{p}})$.

Theorem 3. Suppose that Conditions $[B1]_4$, $[B2]_4$, $[B3]$, $[B4]$ and $[B5]$ are fulfilled. Then

(a) $\sqrt{n}(\tilde{\theta}_n - \theta^*) \longrightarrow^{d_s} \Gamma^{-1/2}\zeta$ **as** $n \to \infty$.

(b) $E\big[f\big(\sqrt{n}(\tilde{\theta}_n - \theta^*)\big)\big] \longrightarrow \mathbb{E}\big[f\big(\Gamma^{-1/2}\zeta\big)\big]$ **as** $n \to \infty$ **for all** $f \in C_{\uparrow}(\mathbb{R}^{\mathsf{p}})$.

Example: A point process driven by a diffusion process

$$\lambda^n(t,\theta) = \lambda^\infty(t,\theta) = g(V_t,\theta)$$

for $t \in I$.

- V_t: a non-degenerate multi-dimensional diffusion process as the covariate
- For the non-degeneracy of χ_0, there are
  - an analytic criterion
  - a geometric criterion.
  - cf. Uchida-Y (SPA 2013)

Support function

Let

$$Q(x,\theta,\theta^*) = g(x,\theta)^{-1}g(x,\theta^*) - 1 - \log\left(g(x,\theta)^{-1}g(x,\theta^*)\right)$$

then

$$-2\mathbb{Y}(\theta) = \frac{1}{T}\int_0^T Q(V_t,\theta,\theta^*)g(V_t,\theta^*)dt.$$

A support function f is a function such that

$$Q(x,\theta,\theta^*)g(x,\theta^*)|\theta-\theta^*|^{-2} \geq |f(x,\theta)|^{\varrho},$$

for a constant $\varrho \in (0,\infty)$. Recall

$$\chi_0 = \inf_{\theta\neq\theta^*}\frac{-\mathbb{Y}(\theta)}{|\theta-\theta^*|^2} \geq \inf_{\theta\neq\theta^*}\frac{1}{2T}\int_0^T |f(V_t,\theta)|^{\varrho}dt.$$

Analytic criterion: nondegeneracy of a tensor field

- For simplicity, let $\mathsf{d} = 1$ and suppose that V is a nondegenerate Itô process.
- Suppose that $\mathcal{X}_0$ is a neighborhood of compact supp$\mathcal{L}\{V_0\}$, and that Θ is compact.
- For each $(x_0, \theta) \in \mathcal{X}_0 \times \Theta$, $\max_{j=0,...,J-1} \left|\partial_x^j f(x_0, \theta)\right| > 0$.

Then [B5] holds.

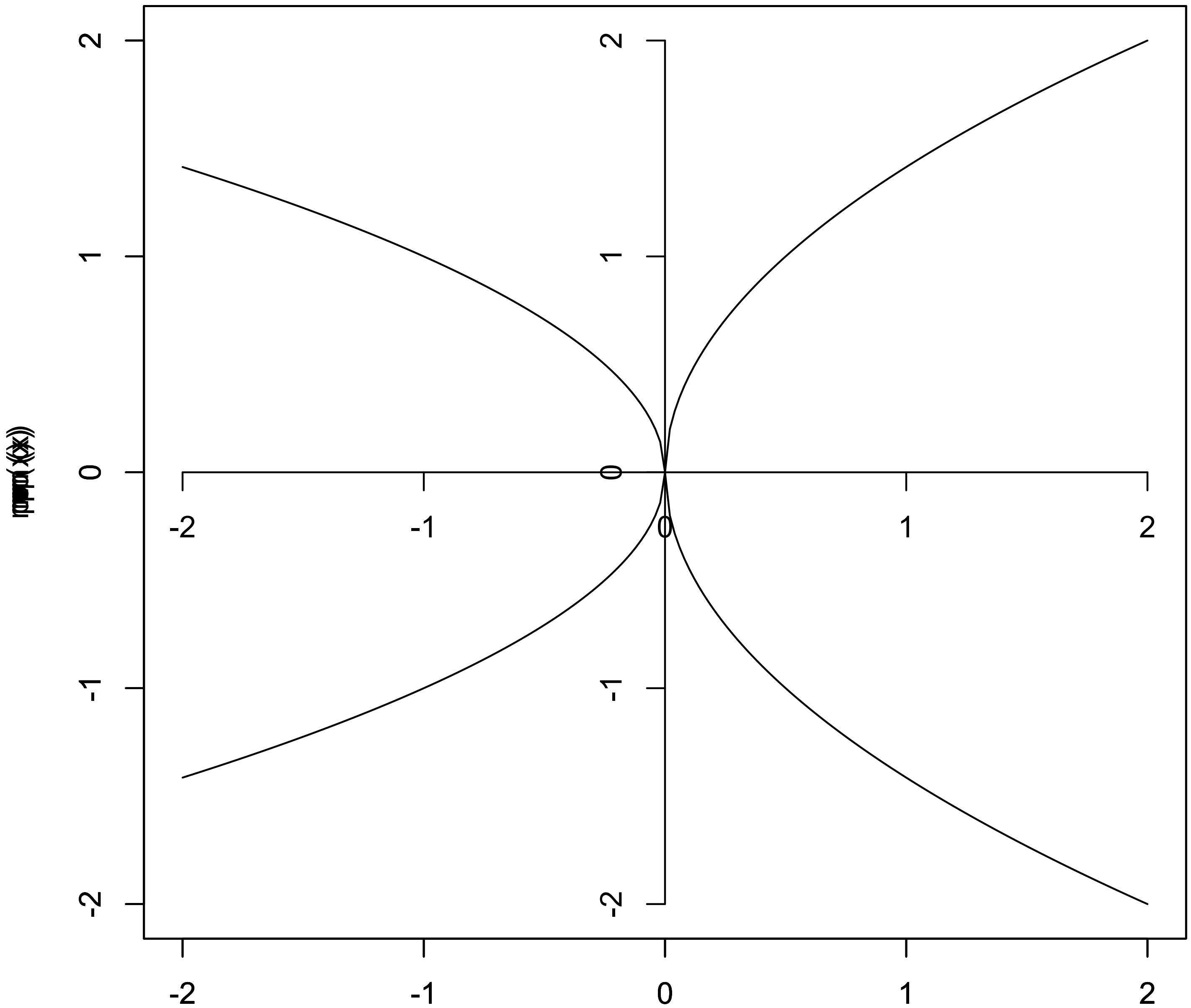
- Remarks.
  - Similar condition in the multi-dimensional case.
  - It is possible to give a condition for a degenerate diffusion on manifold. However the condition becomes much more complicated. (Uchida and Y LeMans2009, ISM RM2011, Paris2012)

Geometric criterion

- **Example.**
  - $f(x,\theta) = x_1 x_2 (x_1 - \theta_1 x_2^2)(\theta_2 x_1 + x_2^2)$
  - $V = (V_t) = (V_{1,t}, V_{2,t})$: **a nondegenerate diffusion with uniform initial distribution on** $\mathrm{supp}\mathcal{L}\{V_0\} = \{0\} \times [0,1]$.
  - **Show**

$$P\left[\inf_{\theta} \int_0^1 |f(V_t,\theta)|^2 dt < \frac{1}{r}\right] \leq \frac{C_L}{r^L}.$$

  - **The null set** $\{f = 0\}$ **is not a regular submanifold.**


-2
-1
0
1
2

Geometric criterion

[A3′] supp$\mathcal{L}\{V_0\}$ is compact, there exists a function $f : U \times \Theta \longrightarrow \mathbb{R}$ for some open neighborhood U of supp$\mathcal{L}\{X_0\}$ and the following conditions are satisfied.

(i) For some $\varrho \in (0, \infty)$, $Q(x, \theta, \theta^*)|\theta - \theta^*|^{-2} \geq |f(x, \theta)|^{\varrho}$ for all $(x, \theta) \in U \times (\Theta \setminus \{\theta^*\})$.

(ii) For each $x_0 \in U$, there exist a neighborhood B in U of x_0 and a covering $\{\Theta_k\}_{k=1,...,\bar{k}}$ of Θ such that for each $k = 1, ..., \bar{k}$, there exist $\xi_0 \in \mathbb{S}$, $J \in \mathbb{N}$, some positive numbers M, c, ϵ_0, K_j $(j = 1, ..., J)$ and some functions $\Psi_j : P^{\perp}_{\xi_0} B \times \Theta_k \longrightarrow \mathbb{R}$ such that

(a) each function $P^{\perp}_{\xi_0} B \ni y \mapsto \Psi_j(y, \theta) \in \mathbb{R}$ is M-Lipschitz continuous for all $\theta \in \Theta_k$,

(b) for $(x,\theta) \in B \times \Theta_k$,

$$|f(x,\theta)| \geq c \prod_{j=1}^{J} \left(|\xi_0 \cdot x - \Psi_j(P_{\xi_0}^{\perp} x, \theta)| \wedge \epsilon_0\right)^{K_j}.$$

Remarks

- In $[A3']$, $\bar{k}$ may depend on x_0.
- The null set

$$\{x \in B;\ f(x,\theta) = 0\} \subset \bigcup_{j=1}^{J} \{x \in B;\ \xi_0 \cdot x = \Psi_j(P_{\xi_0}^{\perp} x, \theta)\}$$

under $[A3']$(ii), that is, the graph of the functions Ψ_j covers locally the null set of f.

Theorem 4. (Uchida-Y arXiv2012 , SPA2013)

$[A3']$ + nondegenerate Itô process $V \Rightarrow [B5]$,

and hence QLA.

Example: A lookback regression

$$\lambda^n(t,\theta) = \lambda^\infty(t,\theta) \; = \; g(t,\gamma) + \int_0^t a e^{-b(t-s)} V_s ds$$

for $t \in I$.

- V_s: a positive non-degenerate diffusion process
- $g(t,\gamma)$: a polynomial taking non-negative values on the interval I
- (γ, a, b): unknown parameters
- The non-degeneracy of χ_0 is not trivial but provable. The exponential kernel is not essential.
- A multi-dimensional extension is possible.

Example: A non-stationary Hawkes process

- **The parametric model of two-dimensional Hawekes process with intensity process**

$$\lambda^n(t,\theta) = g(t,\gamma) + \int_{\hat{T}_0}^{t-} e^{-b(t-s)} A n^{-1} dN_s^n \quad (5)$$

with $\theta = (\gamma, b, A)$.

- **$g_t = g(t,\gamma)$ is an $\mathbb{R}^2$-valued polynomial in t.(inter-day trend)**
- **The non-degeneracy of χ_0 can be proved. Ogihara-Y (arXiv 2015)**

References

[1] **Yoshida, N. (2011)** *Polynomial type large deviation inequalities and quasi-likelihood analysis for stochastic differential equations*, **Annals of the Institute of Statistical Mathematics 63, 431–479**

[2] **Ogihara, T, Yoshida, N. (2015)** *Quasi likelihood analysis of point processes for ultra high frequency data*, **arXiv:1512.01619**

[3] **Clinet, S., Yoshida, N. (2015)** *Statistical Inference for Ergodic Point Processes and Limit Order Book* **arXiv:1512.01899**

[4] **Muni Toke, I., Yoshida, N. (2016)** *Modelling intensities of order flows in a limit order book* **arXiv:1602.03944**

三春滝桜

2011年 春