Estimating the inter-arrival time density of Markov renewal processes under structural assumptions on the transition distribution

Priscilla E. Greenwood
School of Mathematics and Statistics
and Mathematical and Computational Sciences Modeling Center
Arizona State University
Tempe, AZ 85287-1804, USA
pgreenw@math.la.asu.edu

Anton Schick *
Department of Mathematical Sciences
Binghamton University
Binghamton, NY 13902-6000, USA
anton@math.binghamton.edu

Wolfgang Wefelmeyer
Mathematical Institute
University of Cologne
50931 Cologne, Germany
wefelm@math.uni-koeln.de

Abstract

We consider a stationary Markov renewal process whose inter-arrival time density depends multiplicatively on the distance between past and present state of the embedded chain. This is appropriate when the jump size is governed by influences that accumulate over time. Then we can construct an estimator for the inter-arrival time density that has the parametric rate of convergence. The estimator is a local von Mises statistic. The result carries over to the corresponding semi-Markov process.

AMS 2000 subject classifications. Primary: 62G07, 62M05; secondary: 62G20.

Keywords. Local von Mises statistic, asymptotically linear estimator, density estimator.

1 Introduction

Let $(X_0, T_0), \dots, (X_n, T_n)$ be observations of a Markov renewal process with real state space. Assume that the embedded Markov chain $X_0, X_1, \dots$ has transition density $q(x, y)$ and is stationary. Then X_j has a stationary density, say p_1, and the stationary density of

*The research of Anton Schick was supported in part by NSF Grant DMS0906551

(X_{j-1}, X_j) at (x, y) is $p_2(x, y) = p_1(x)q(x, y)$. Write $q^{(m)}(x, y)$ for the m-step transition density, and write $V_j = T_j - T_{j-1}$ for the inter-arrival times. Assume that V_j has a conditional density $r(x, y, v)$ at v given $(X_{j-1}, X_j) = (x, y)$. Then the Markov chain $(X_1, V_1), (X_2, V_2), \dots$ has a transition density from $(X_{j-1}, V_{j-1}) = (x, u)$ to $(X_j, V_j) = (y, v)$ that factors as $s(x; y, v) = q(x, y)r(x, y, v)$ and does not depend on u.

We are interested in estimating the stationary density of V_j at $v > 0$,

$$\varrho(v) = \iint p_2(x, y)r(x, y, v)\, dx\, dy.$$

Write $f_s(x) = f(x/s)/s$ for a density f scaled by $s > 0$. Let k be a kernel and $b = b_n$ a bandwidth that tends to zero as n goes to infinity. The usual estimator of $\varrho(v)$ is a kernel estimator

$$\tilde{\varrho}(v) = \frac{1}{n}\sum_{j=1}^{n} k_b(v - V_j).$$

Kernel estimators have been studied under various mixing conditions; recent papers are Liebscher (1992), Lu (2001), Prieur (2001), and Ragache and Wintenberger (2006). The convergence rates are similar or worse than those for the case of i.i.d. observations and depend on the degree of smoothness of ϱ and on the choice of bandwidth.

It is the purpose of this paper to show that, under additional structural assumptions on r, we can construct an estimator for ϱ that converges at the faster, parametric, rate $n^{-1/2}$. Specifically, we assume that the inter-arrival times depend *multiplicatively* on the distance between the past and present states X_{j-1} and X_j of the embedded Markov chain,

$$V_j = |X_j - X_{j-1}|^{\alpha} W_j,$$

where $\alpha > 0$ is known and the W_j are positive, independent with common density g, and independent of the embedded Markov chain. This means that the jump sizes are roughly proportional to a power of the inter-arrival times.

Such a model is plausible when the jumps result from influences that accumulate over time. An example is the size of a tectonic earthquake. Here the stress builds up over time because it is due to the relative motion of adjacent tectonic plates. See Shimazaki and Nakata (1980), Murray and Segall (2002), and Corral (2006), who also considers the *spatial* distribution of earthquakes. Similarly, the number of susceptibles will rise again after an epidemic, due to loss of immunity in the population, so the size of the next epidemic will be larger if more time has elapsed. It is interesting that in finance, shorter waiting times imply larger jumps, because higher volatility leads to faster trading. This would correspond to $\alpha < 0$ in our model and could be treated similarly. See Raberto et al. (2002) and Meerschaert and Scalas (2006).

Since $Z_j = |X_j - X_{j-1}|^{\alpha}$ and $W_j = |X_j - X_{j-1}|^{-\alpha} V_j$ are independent, $Z_i W_j$ is distributed as $Z_j W_j$, and we can estimate the density of $V_j = Z_j W_j$ at $v > 0$ by a *local von Mises* *statistic*

$$(1.1)\qquad \hat\varrho(v) = \frac{1}{n^2}\sum_{i=1}^n\sum_{j=1}^n k_b(v - Z_iW_j) = \frac{1}{n^2}\sum_{i=1}^n\sum_{j=1}^n k_b\Big(v - |X_i - X_{i-1}|^\alpha \frac{T_j - T_{j-1}}{|X_j - X_{j-1}|^\alpha}\Big).$$

The von Mises statistic is called *local* because its kernel involves a scale parameter b that tends to zero. For i.i.d. observations $X_1, \dots, X_n$, similar local U-statistics for densities of transformations $a(X_1, \dots, X_m)$ with $m \geq 2$ have been studied by Frees (1994) and Giné and Mason (2007), among others. For (non-local) von Mises statistics based on *dependent* observations we refer to Dehling (2006). The von Mises statistic (1.1) considered here is both local and based on dependent observations. In Section 2 we give conditions under which the density estimator $\hat\varrho(v)$ is asymptotically linear and asymptotically normal and therefore converges to $\varrho(v)$ at the parametric rate $n^{-1/2}$.

Consider the semi-Markov process Z_t, $t \geq 0$, corresponding to the above Markov renewal process. Suppose we observe a path Z_t, $0 \leq t \leq n$. Set $N = \max\{j : T_j \leq n\}$. The result of Section 2 carries over to the semi-Markov process by replacing the sample size n by the random sample size N.

2 Result

Assume that the inter-arrival times are of the form $V_j = T_j - T_{j-1} = |X_j - X_{j-1}|^\alpha W_j$, where α is a known positive constant and the non-negative random variables $W_1, W_2, \dots$ are independent and identically distributed with common density g and are independent of the embedded Markov chain. Recall that q and $q^{(m)}$ denote the one-step and m-step transition densities of this chain.

We call the embedded Markov chain *uniformly ergodic* if

$$\sup_{x\in\mathbb{R}}\sup_{|f|\leq 1}\Big|\int (q^{(m)}(x,y) - p_1(y))f(y)\,dy\Big| \to 0 \quad \text{as } m \to \infty.$$

This implies uniform ergodicity at a geometric rate; see e.g. Meyn and Tweedie (1993, Chapter 16).

The conditional density at $y > 0$ of $|X_j - X_{j-1}|$ given $X_{j-1} = x$ is

$$\gamma(x,y) = q(x, x+y) + q(x, x-y).$$

Setting $Z_j = |X_j - X_{j-1}|^\alpha$ and $\beta = 1/\alpha$, the conditional density at $y > 0$ of Z_j given $X_{j-1} = x$ is

$$\zeta(x,y) = \beta y^{\beta-1}\gamma(x, y^\beta).$$

Hence the stationary density at $y > 0$ of Z_j is

$$h(y) = \int p_1(x)\zeta(x,y)\,dx = \beta y^{\beta-1}\int p_1(x)\gamma(x,y^\beta)\,dx.$$

We can therefore write the stationary density of $V_j = Z_j W_j$ at some positive v as a scale mixture in two different ways,

$$\varrho(v) = \int h_w(v) g(w)\, dw = \int g_y(v) h(y)\, dy.$$

See also Rohatgi and Ehsanes Saleh (2001), Section 4.4, or Glen et al. (2004). Of course $\gamma(x,y)$, $\zeta(x,y)$, $h(y)$ and $\varrho(y)$ equal zero if y is not positive.

Recall that a function f from $\mathbb{R}$ to $\mathbb{R}$ is *Hölder at t with exponent ξ* if $|f(s) - f(t)| \le L|s-t|^\xi$ for some constant L and all s in a neighborhood of t.

Theorem 1. *Suppose the embedded Markov chain is uniformly ergodic. Let the functions $v \mapsto vg(v)$ and $v \mapsto vh(v)$ be bounded and continuous. Let ϱ be Hölder at v with exponent $\xi > 1/2$. Assume that the kernel k is a bounded symmetric density with support $[-1,1]$ and that the bandwidth $b = b_n$ fulfills $nb_n \to \infty$ and $nb_n^{2\alpha} \to 0$. Then*

$$\hat\varrho(v) = \varrho(v) + \frac{1}{n}\sum_{j=1}^n \big(h_{W_j}(v) - \varrho(v) + g_{Z_j}(v) - \varrho(v)\big) + o_P(n^{-1/2}).$$

Proof. First we calculate the Hoeffding decomposition of $\hat\varrho(v)$. Since Z_i and W_j are independent, the conditional expectation of $k_b(v - Z_i W_j)$ given $W_j = w > 0$ is given by

$$\begin{aligned} H(w) = E(k_b(v - Z_i W_j)|W_j = w) &= \int k_b(v - zw) h(z)\, dz \\ &= \int k_b(v-z) h_w(z)\, dz = \int h_w(v - bu) k(u)\, du. \end{aligned}$$

Similarly, the conditional expectation of $k_b(v - Z_i W_j)$ given $Z_i = z > 0$ is

$$G(z) = E(k_b(v - Z_i W_j)|Z_j = z) = \int k_b(v - zw) g(w)\, dw = \int g_z(v - bu) k(u)\, du.$$

The expectation of $k_b(v - Z_i W_j)$ is

$$R = E[H(W_1)] = E[G(Z_1)] = E[k_b(v - Z_1 W_1)] = \int \varrho(v - bu) k(u)\, du.$$

Hence the Hoeffding decomposition of $\hat\varrho(v)$ is

$$\hat\varrho(v) = R + \bar G_n + \bar H_n + U_n$$

with

$$\bar G_n = \frac{1}{n}\sum_{j=1}^n \big(G(Z_j) - E[G(Z_1)]\big), \quad \bar H_n = \frac{1}{n}\sum_{j=1}^n \big(H(W_j) - E[H(W_1)]\big),$$

and

$$U_n = \frac{1}{n^2}\sum_{i=1}^n \sum_{j=1}^n \big(k_b(v - Z_i W_j) - G(Z_i) - H(W_j) + R\big).$$

To show that $U_n = o_P(n^{-1/2})$, we introduce the following notation,

$$U_{ij} = K(Z_i, W_j) = k_b(v - Z_i W_j) - G(Z_i) - H(W_j) + R.$$

Note that $W_1, \dots, W_n$ are independent with a density g and independent of the Markov chain $X_0, \dots, X_n$, and therefore independent of $Z_i = |X_i - X_{i-1}|^\alpha$ for $i = 1, \dots, n$. Hence two summands U_{ij} and U_{kl} are uncorrelated if $j \neq l$. From this we immediately derive that

$$\begin{aligned} nE[U_n^2] &= \frac{1}{n^3} \sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n E[U_{ij} U_{kj}] \\ &= \frac{1}{n^2} \sum_{i=1}^n \sum_{k=1}^n \int E[K(Z_i, w) K(Z_k, w)] g(w)\, dw. \end{aligned}$$

This shows that $nE[U_n^2] = E[A_n^2]$, where

$$A_n = \frac{1}{n} \sum_{j=1}^n K(Z_j, W_1).$$

Now set

$$K_0(X_{j-1}, W_1) = E(K(Z_j, W_1) | X_{j-1}, W_1) = \int q(X_{j-1}, y) K(|y - X_{j-1}|^\alpha, W_1)\, dy.$$

Let

$$A_{n0} = \frac{1}{n} \sum_{j=1}^n K_0(X_{j-1}, W_1).$$

Then $n(A_n - A_{n0}) = \sum_{j=1}^n \big(K(Z_j, W_1) - K_0(X_{j-1}, W_1)\big)$ is a martingale, and

$$E[(A_n - A_{n0})^2] = n^{-1} E[(K(Z_1, W_1) - K_0(X_0, W_1))^2] \leq n^{-1} E[K^2(Z_1, W_1)] = O((nb)^{-1}),$$

where for the last equality we have used

$$E[K^2(Z_1, W_1)] \leq E[k_b^2(v - Z_1 W_1)] = k_b^2 * \rho(v) = b^{-1} \int \rho(v - bu) k^2(u)\, du = O(b^{-1}).$$

Next we have

$$\begin{aligned} E[A_{n0}^2] &= \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n E[K_0(X_{i-1}, W_1) K_0(X_{j-1}, W_1)] \\ &= \frac{1}{n} E[K_0^2(X_0, W_1)] + \frac{2}{n^2} \sum_{m=1}^{n-1} (n - m) \int E[K_0(X_0, w) K_0(X_m, w)] g(w)\, dw. \end{aligned}$$

The embedded Markov chain is uniformly ergodic and therefore geometrically uniformly ergodic. Hence there exist constants $a > 0$ and $0 < \lambda < 1$ such that

$$D_m = \sup_{x \in \mathbb{R}} \sup_{|f| \leq 1} \Big| \int (q^{(m)}(x, y) - p_1(y)) f(y)\, dy \Big| \leq a \lambda^m.$$

Let B be a bound for the kernel k. Then K_0 is bounded by $4Bb^{-1}$. For positive w, we have $\int K_0(x,w)p_1(x)\,dx = 0$ and thus

$$|E[K_0(X_0,w)K_0(X_m,w)]| = \Big| \int K_0(x,w) \int q^m(x,y)K_0(y,w)\,dy\,p_1(x)\,dx \Big|$$
$$\leq 4Bb^{-1}D_m \int p_1(x)|K_0(x,w)|\,dx.$$

From this we immediately derive

$$\frac{1}{n^2}\sum_{m=1}^{n-1}(n-m)\int |E[K_0(X_0,w)K_0(X_m,w)]|g(w)\,dw \leq \frac{1}{n}\sum_{m=1}^{n-1} 4Bb^{-1}a\lambda^m E[|K_0(X_0,W_1)|].$$

This and the inequality $E[|K_0(X_0,W_1)|] \leq 4R$ show that

$$nE[U_n^2] \leq 2E[(A_n - A_{n0})^2] + 2E[A_{n0}^2] = O((nb)^{-1}).$$

We obtain that the centered local von Mises statistic $\hat{\varrho}(v) - R$ is approximated by a sum of two smoothed and centered empirical estimators,

$$n^{1/2}(\hat{\varrho}(v) - R) = n^{-1/2}\sum_{j=1}^{n}\big(G(Z_j) - E[G(Z_1)] + H(W_j) - E[H(W_1)]\big) + o_P(1).$$

The kernel k has bounded support, $\int k(u)\,du$ equals 1 and $\int |u|^\xi k(u)\,du$ is finite. Using this, the Hölder property of ϱ, and the properties of the bandwidth, we obtain

$$|R - \varrho(v)| \leq \Big| \int (\rho(v - bu) - \rho(v))k(u)\,du \Big| \leq Lb^\xi \int |u|^\xi k(u)\,du = O(b^\xi) = o(n^{-1/2}).$$

Let $\chi(v) = vh(v)$. Then $h_w(v) = \chi(v/w)/v$. For $v, w > 0$ and $\delta < v/2$ we have

$$|h_w(v - \delta) - h_w(v)| \leq \frac{3}{v}\sup_{x>0}\chi(x).$$

Furthermore, by continuity of χ,

$$h_w(v-\delta) \to h_w(v) \quad \text{as} \quad \delta \to 0.$$

From the dominated convergence theorem we therefore obtain

$$E[(H(W_1) - h_{W_1}(v))^2] \leq \iint \big(h_w(v - bu) - h_w(v)\big)^2 k(u)\,du\,g(w)\,dw \to 0.$$

An analogous result holds for G in place of H. It follows that the two smoothed empirical estimators are approximated by their unsmoothed versions,

$$n^{-1/2}\sum_{j=1}^{n}(G(Z_j) - E[G(Z_1)]) = n^{-1/2}\sum_{j=1}^{n}\big(g_{Z_j}(v) - \varrho(v)\big) + o_P(1),$$
$$n^{-1/2}\sum_{j=1}^{n}(H(W_j) - E[H(W_1)]) = n^{-1/2}\sum_{j=1}^{n}\big(h_{W_j}(v) - \varrho(v)\big) + o_P(1).$$

The assertion follows. □

The requirements on the bandwidth b are satisfied by the choice $b = b_n = \log n/n$, even if the value of ξ is unknown. Larger values are possible for b if the value of ξ is known. For example, if $\xi = 1$, then we can take $b = b_n = (n \log n)^{-1/2}$. Even larger values of b are possible if we impose additional smoothness assumptions on ϱ. For example, if ϱ is differentiable in a neighborhood of v and its derivative is Hölder at v with exponent ξ, then the bias will be of order $O(b^{1+\xi})$ and we can relax $nb^{2\xi} \to 0$ to $nb^{2+2\xi} \to 0$.

The stationary density ϱ of the inter-arrival time $V_j = T_j - T_{j-1}$ is Hölder at $v > 0$ with exponent ξ if $\int y^{-1-\xi} h(y)\, dy$ is finite and the density g of W_j satisfies

$$|g(s) - g(t)| \le L|s - t|^{\xi}, \quad 0 < s < t,$$

for some constant L. This follows from the inequality

$$|\varrho(u) - \varrho(v)| \le \int h(y) y^{-1} |g(u/y) - g(v/y)|\, dy \le L|u - v| \int y^{-1-\xi} h(y)\, dy.$$

A similar sufficient condition can be given by switching the roles of h and g.

Theorem 1 implies that $\hat{\varrho}(v)$ is $n^{1/2}$-consistent. More precisely, $n^{1/2}(\hat{\varrho}(v) - \varrho(v))$ is asymptotically normal by a central limit theorem for Markov chains. Since the W_i are independent of the Z_j, the asymptotic variance of $\hat{\varrho}(v)$ is the sum of the asymptotic variances of the two terms $n^{-1/2} \sum_{j=1}^n \big(h_{W_j}(v) - \varrho(v)\big)$ and $n^{-1/2} \sum_{j=1}^n \big(g_{Z_j}(v) - \varrho(v)\big)$. Since the W_j are independent, the asymptotic variance of the first term is $E[h^2_{W_1}(v)] - \varrho^2(v)$. The asymptotic variance of the second term is

$$E[g^2_{Z_1}(v)] - \varrho^2(v) + 2 \sum_{s=2}^{\infty} E[(g_{Z_1}(v) - \varrho(v)) g_{Z_s}(v)].$$

We can however not expect *functional* results for the process $v \mapsto n^{1/2}(\hat{\varrho}(v) - \varrho(v))$, in general. This will be shown elsewhere. Comparable non-regular behavior is shown by local U-statistics for the density of $|X_1|^\alpha + |X_2|^\alpha$ when $\alpha \ge 2$; see Schick and Wefelmeyer (2009a) and (2009b).

One can show that the local von Mises statistic $\hat{\varrho}(v)$ is asymptotically efficient unless we know more about the transition distribution of the Markov renewal process. We now discuss briefly improvements over $\hat{\varrho}(v)$ in two cases.

1. Assume that we have a parametric model for the density g of the W_j. For example, let the W_j be *exponentially distributed* with unknown mean $\mu > 0$, i.e., $g(w) = \exp(-w/\mu)/\mu$ for $w > 0$. Then we can estimate $\varrho(v)$ as follows. Estimate μ by $\hat{\mu} = (1/n) \sum_{j=1}^n W_j$. Estimate the stationary density h of Z_j by a kernel estimator

$$\hat{h}(y) = \frac{1}{n} \sum_{j=1}^n k_b(y - Z_j).$$

Then use the representation $\varrho(v) = \int h_w(v)g(w)\,dw$ to estimate $\varrho(v)$ by the plug-in estimator

$$\hat{\varrho}(v) = \frac{1}{\hat{\mu}} \int_0^\infty \hat{h}_w(v) e^{-w/\hat{\mu}}\,dw.$$

2. Assume instead that we know more about the structure of the transition density q of the embedded Markov chain. For example, assume that this chain is *autoregressive*, say $X_j = \vartheta X_{j-1} + \varepsilon_j$, where $|\vartheta| < 1$ and the innovations ε_j are i.i.d. with mean zero, finite variance, and positive density f. Then $q(x, y) = f(y - \vartheta x)$, and we can replace the kernel estimator for the stationary density of Z_j by a plug-in estimator

$$\hat{h}(y) = \beta y^{\beta-1} \int \hat{p}_1(x)\hat{\gamma}(x, y^\beta)\,dx$$

with

$$\hat{\gamma}(x, y) = \big(\hat{f}(x + y - \hat{\vartheta}x) + \hat{f}(x - y - \hat{\vartheta}x)\big)\mathbf{1}(y > 0).$$

Here we can use the least squares estimator $\hat{\vartheta}$ to estimate ϑ, and a kernel estimator $\hat{f}$ for f based on residuals $\hat{\varepsilon}_j = X_j - \hat{\vartheta}X_{j-1}$. The simplest choice for $\hat{p}_1$ is a kernel estimator based on observations based on the embedded Markov chain. Improvements over such kernel estimators are studied in Schick and Wefelmeyer (2007, 2008, 2009c).

Acknowledgments

The authors wish to thank the referee for several suggestions that improved the paper.

References

[1] Corral, A. (2006). Universal earthquake-occurrence jumps, correlations with time, and anomalous diffusion. *Phys. Rev. Lett.* **97**, 178501, 4 pp.

[2] Dehling, H. (2006). Limit theorems for dependent U-statistics. In: *Dependence in Probability and Statistics*, (P. Bertail, P. Doukhan and P. Soulier, eds.), 65–86, Lecture Notes in Statistics 187, Springer, New York.

[3] Frees, E. W. (1994). Estimating densities of functions of observations. *J. Amer. Statist. Assoc.* **89**, 517–525.

[4] Giné, E. and Mason, D. (2007). On local U-statistic processes and the estimation of densities of functions of several variables. *Ann. Statist.* **35**, 1104–1145.

[5] Glen, A. G., Leemis, L. M. and Drew, J. H. (2004). Computing the distribution of the product of two continuous random variables. *Comput. Statist. Data Anal.* **44**, 451–464.

[6] Liebscher, E. (1992). Density estimation for Markov chains. *Statistics* **23**, 27–48.

[7] Lu, Z. (2001). Asymptotic normality of kernel density estimators under dependence. *Ann. Inst. Statist. Math.* **53**, 447–468.

[8] Meerschaert, M. M. and Scalas, E. (2006). Coupled continuous time random walks in finance. *Phys. A* **370**, 114–118.

[9] Meyn, S. P. and Tweedie, R. L. (1993). *Markov Chains and Stochastic Stability.* Communications and Control Engineering Series, Springer, London.

[10] Murray, J. and Segall, P. (2002). Testing time-predictable earthquake recurrence by direct measurement of strain accumulation and release. *Nature* **419**, 287–291.

[11] Prieur, C. (2001). Density estimation for one-dimensional dynamical systems. *ESAIM Probab. Statist.* **5**, 51–76.

[12] Raberto, M., Scalas, E. and Mainardi, F. (2002). Waiting-times and returns in high-frequency financial data: an empirical study. *Phys. A* **314**, 749–755.

[13] Ragache, N. and Wintenberger, O. (2006). Convergence rates for density estimators of weakly dependent time series. In: *Dependence in Probability and Statistics* (P. Bertail, P. Doukhan and P. Soulier, eds.), 349–372, Lecture Notes in Statistics 187, Springer, New York.

[14] Rohatgi, V. K. and Ehsanes Saleh, A. K. Md. (2001). *An Introduction to Probability and Statistics. Second Edition.* Wiley Series in Probability and Statistics, Wiley, New York.

[15] Schick, A. and Wefelmeyer, W. (2007). Uniformly root-n consistent density estimators for weakly dependent invertible linear processes. *Ann. Statist.* **35**, 815–843.

[16] Schick, A. and Wefelmeyer, W. (2008). Root-n consistency in weighted L_1-spaces for density estimators of invertible linear processes. *Stat. Inference Stoch. Process.* **11**, 281–310.

[17] Schick, A. and Wefelmeyer, W. (2009a). Convergence rates of density estimators for sums of powers of observations. *Metrika* **69**, 249–264.

[18] Schick, A. and Wefelmeyer, W. (2009b). Non-standard behavior of density estimators for sums of squared observations. *Statist. Decisions* **27**, 55–73.

[19] Schick, A. and Wefelmeyer, W. (2009c). Improved density estimators for invertible linear processes. *Comm. Statist. Theory Methods* **38**, 3123–3147.

[20] Shimazaki, K. and Nakata, T. (1980). Time-predictable recurrence model for large earthquakes. *Geophys. Res. Lett.* **7**, 279-282.