Ann. Inst. Statist. Math.
Vol. 42, No. 1, 89–98 (1990)

# PARAMETER ESTIMATION FOR THE SIMPLE SELF-CORRECTING POINT PROCESS

NOBUO INAGAKI[1*] AND TOSHIHARU HAYASHI[2]

[1]*Department of Mathematical Science, Faculty of Engineering Science, Osaka University, Toyonaka, Osaka 560, Japan*
[2]*Department of Mathematics, College of Integrated Arts and Sciences, University of Osaka Prefecture, Sakai, Osaka 591, Japan*

(Received February 10, 1988; revised May 18, 1989)

**Abstract.** We consider the simple self-correcting point process whose intensity takes only the two levels, $a$ and $b$, where the level $a$ $(0 < a < 1)$ is the state of the intensity for low stress, and the level $b$ $(1 < b < \infty)$ is for high stress. Then, the maximum likelihood estimators of $a$ and $b$ and their asymptotic distributions are explicitly shown. These results may be instructive and suggestive for studying more general cases of self-correcting point processes.

*Key words and phrases*: Self-correcting point process, stress release process, maximum likelihood estimator, invariant distribution, transition probability, ergodicity, local asymptotic normality (LAN).

## 1. Self-correcting point process

We consider a point process $N(\cdot)$ on the time $[0,\infty)$. Let $\tau_1, \tau_2, \ldots$ be the occurrence times of the events, and let $N(t) = \#\{\tau_i\colon 0 \le \tau_i \le t\}$ be the number of the occurrence points for the time $[0, t]$ with $N(0) = 0$. Let us denote the $\sigma$-field generated by the events for the time interval $[0, t]$ by $\mathscr{F}_t = \sigma(\{N(s)\colon 0 \le s \le t\})$. Suppose $N(t)$ is a pure birth process with the instantaneous birth rate at time $t$ given $N(t) = n$:

$$\lambda(t) = \lambda(t|\mathscr{F}_{t-}) = \rho\phi(\rho t - n)\,, \tag{1.1}$$

where $\rho$ is a positive constant and $\phi$ satisfies the following conditions:

(SC1) $0 \le \phi(x) < \infty$, for any $x \in R$.

(SC2) There is a positive constant $c > 0$ such that $\phi(x) \ge c$, for every $x > 0$.

*Work supported by the Institute of Statistical Mathematics (63-kyouken-4).

(SC3) (i) $\liminf_{x \to \infty} \phi(x) > 1$,

(ii) $\limsup_{x \to -\infty} \phi(x) < 1$.

The last condition (SC3) implies that a larger positive value of $x = \rho t - N(t)$, caused by fewer occurrence points in the time interval $[0, t]$, makes the intensity $\phi(x)$ larger and results in more occurrence points in the forthcoming time interval. Conversely, more occurrence points cause a smaller negative value of $x = \rho t - N(t)$, which makes the intensity $\phi(x)$ smaller and results in fewer occurrence points in the forthcoming time interval. That is, the point process is self-correcting.

Isham and Westcott (1979) introduced such a point process and called it a self-correcting point (SCP) process. By changing the scale of time, we may take $\rho = 1$ without loss of generality. They showed the following results about the mean and variance of the SCP process $N(t)$.

THEOREM 1.1. (*Isham and Westcott* (1979))
(i) $\limsup_{t \to \infty} |\mu(t) - t| < \infty$.
(ii) $\limsup_{t \to \infty} |V(t)| < \infty$.

Vere-Jones and Ogata (1984) considered the SCP process with the conditional intensity of exponential type:

$$\lambda(t) = \exp\{\alpha + \beta(t - N(t))\},$$

and obtained the ergodicity of $X(t) = t - N(t)$:

THEOREM 1.2. *For a measurable function $h(\cdot)$ which is exponentially dominated, it holds that*

$$\text{(1.2)} \qquad \text{p-}\lim_{T \to \infty} T^{-1} \int_0^T h(X(t))dt = \sum_{j=-\infty}^{\infty} \pi_j E\left[\int_0^1 h(X(t))dt \,|\, X(0) = j\right],$$

*where $\{\pi_j\}_{j=-\infty}^{\infty}$ is the invariant distribution of the skeleton Markov chain $\{X(n)\}_{n=0}^{\infty}$.*

Furthermore, Ogata and Vere-Jones (1984) showed the asymptotic normality of the maximum likelihood estimators of the parameters $\alpha$ and $\beta$. Vere-Jones (1988) called $X(t) = t - N(t)$ the stress release process and investigated its moments.

Hayashi (1986) extended the results due to Vere-Jones and Ogata to the general self-correcting point process which is subject to the previous conditions (SC1)–(SC3) and the additional one:

(SC4) For any $K > 0$, there is a positive number $M$ such that

$\phi(x) \leq M$, for every $x \leq K$.

Hayashi (1988) obtained the asymptotic normality of the maximum likelihood estimators of the parameters $\alpha$, $\beta$, $\rho$ of the conditional intensity:

$$\lambda(t) = \rho\phi(\beta\{\rho t - N(t-) + \alpha\}) .$$

However, these results could not be written down "explicitly" in terms of the intensity function and its parameters because of the complexity of the intensity function. In the present paper, we consider a very simple conditional intensity defined by

$$\lambda(t) = \rho\phi(\rho t - N(t-)) = \begin{cases} \rho a, & \text{if} \quad \rho t - N(t-) \leq c , \\ \rho b, & \text{if} \quad \rho t - N(t-) > c , \end{cases} \tag{1.3}$$

where $a$, $b$ and $c$ are real constants with $0 < a < 1 < b$. This point process satisfies the conditions (SC1)–(SC4), and we call it a simple self-correcting point (SSCP) process. By changing the location and scale of time, we may take $c = 0$ and $\rho = 1$ without loss of generality. That is, let us consider the SSCP process with the following conditional intensity:

$$\lambda_\theta(t) = \begin{cases} a, & \text{if} \quad X(t-) \leq 0 , \\ b, & \text{if} \quad X(t-) > 0 , \end{cases} \tag{1.4}$$

with the parameter $\theta = (a, b)'$, where $0 < a < 1 < b$ and $X(t) = t - N(t)$, the stress release process. Then, we obtain "explicitly" the maximum likelihood estimator and its asymptotic properties. These results are instructive and suggestive for studying more general cases of the SCP process.

## 2. Likelihood function and maximum likelihood estimator

The likelihood function of a regular point process on the time interval $[0, T]$ is expressed by the conditional intensity as follows (see Snyder (1975)):

$$L_T(\theta) = \exp\left\{\int_0^T \log \lambda(t) dN(t) - \int_0^T \lambda(t) dt\right\} . \tag{2.1}$$

Therefore, the likelihood function of the SSCP process is given by

$$L_T(\theta) = \exp\left[\left\{\int_0^T (\log a)\chi_{A_t}(t-)dN(t) + \int_0^T (\log b)\chi_{B_t}(t-)dN(t)\right\} - \left\{\int_0^T a\chi_{A_t}(t-)dt + \int_0^T b\chi_{B_t}(t-)dt\right\}\right] \tag{2.2}$$

$$= \exp[\{(\log a)N(A_T) + (\log b)N(B_T)\} - \{aL(A_T) - bL(B_T)\}] ,$$

where $\chi_A(t)$ is an indicator function of a set $A$ and

$$\begin{aligned} & A_t = \{s: 0 \le s \le t,\ X(s) \le 0\}, && B_t = \{s: 0 \le s \le t,\ X(s) > 0\} , \\ (2.3)\quad & N(A_T) = \int_0^T \chi_{A_t}(t-)dN(t), && N(B_T) = \int_0^T \chi_{B_t}(t-)dN(t) , \\ & L(A_T) = \int_0^T \chi_{A_t}(t-)dt, && L(B_T) = \int_0^T \chi_{B_t}(t-)dt , \end{aligned}$$

all of which depend on the path. The log-likelihood equations are

$$(\partial/\partial a) \log L_T(\theta) = N(A_T)/a - L(A_T) = 0 ,$$
$$(\partial/\partial b) \log L_T(\theta) = N(B_T)/b - L(B_T) = 0 ,$$

and hence, the maximum likelihood estimators of $a$ and $b$ are

$$(2.4)\qquad \hat{a}_T = N(A_T)/L(A_T) \quad \text{and} \quad \hat{b}_T = N(B_T)/L(B_T) .$$

Let us introduce the process $M(T) = N(T) - \int_0^T \lambda(t)dt$ and define $J(T) = \int_0^T f(t)dM(t)$ as the Lebesgue-Stieltjes integral of the left-continuous and $\mathscr{F}_t$-measurable function $f(t, \omega)$. Then, $J(T)$ has the same properties as the Itô stochastic integral with respect to the Wiener process. Suppose $E\left\{\int_0^T |f(t)|^2\lambda(t)dt\right\} < \infty$, then the following results hold:

THEOREM 2.1. (*Kutoyants* (1984))

(i) $\left\{\int_0^T f(t)dM(t), \mathscr{F}_T: 0 \le T < \infty\right\}$ *is a martingale.*

(ii) $E\left\{\int_0^T f(t)dM(t)\right\} = 0.$

(iii) $E\left[\left\{\int_0^T f(t)dM(t)\right\}^2\right] = E\left\{\int_0^T f(t)^2\lambda(t)dt\right\}.$

The likelihood estimating function is

$$(2.5)\qquad (\partial/\partial\theta) \log L_T(\theta) = \begin{pmatrix} \int_0^T \chi_{A_t}(t-)dM(t)/a \\ \int_0^T \chi_{B_t}(t-)dM(t)/b \end{pmatrix}$$

and so, by the above theorem, we have the Fisher information matrix as follows:

$$I_T(\theta) = \begin{pmatrix} E\{L(A_T)\}/a & 0 \\ 0 & E\{L(B_T)\}/b \end{pmatrix}. \tag{2.6}$$

Furthermore, we have

$$\hat{\theta}_T - \theta = \begin{pmatrix} \hat{a}_T - a \\ \hat{b}_T - b \end{pmatrix} = \begin{pmatrix} \int_0^T \chi_{A_t}(t-)dM(t)/L(A_T) \\ \int_0^T \chi_{B_t}(t-)dM(t)/L(B_T) \end{pmatrix}, \tag{2.7}$$

but it is difficult to obtain the mean vector and covariance matrix of the maximum likelihood estimator.

## 3. Stationary distribution of $\{X(n)\}_{n=0}^{\infty}$

The stress release process $X(t) = t - N(t)$ is an ergodic Markov process and the skeleton $\{X(n)\}_{n=0}^{\infty}$ is an irreducible and aperiodic Markov chain.

### 3.1 *Transition probability*

Let us investigate the transition probability of the skeleton Markov chain: $P_{h,k} = P\{X(n+1) = k \mid X(n) = h\}$. Since $P_{h,k} = 0$ for $k \geq h+2$, the transition probability matrix is in the following form:

$$\begin{pmatrix} \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \cdots & P_{-2,-2} & P_{-2,-1} & 0 & 0 & 0 & 0 & 0 & \cdots \\ \cdots & P_{-1,-2} & P_{-1,-1} & P_{-1,0} & 0 & 0 & 0 & 0 & \cdots \\ \cdots & P_{0,-2} & P_{0,-1} & P_{0,0} & P_{0,1} & 0 & 0 & 0 & \cdots \\ \cdots & P_{1,-2} & P_{1,-1} & P_{1,0} & P_{1,1} & P_{1,2} & 0 & 0 & \cdots \\ \cdots & P_{2,-2} & P_{2,-1} & P_{2,0} & P_{2,1} & P_{2,2} & P_{2,3} & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix}. \tag{3.1}$$

(i) Under the condition given by $X(n) = -h$ $(h > 0)$, $X(n+1)$ can take the values: $X(n+1) = -h+1, -h, -h-1, -h-2, \ldots$ with the probabilities,

$$P_{-h,-k} = P_{-h,-(h-1)-(k-h+1)} = e^{-a}a^{k-h+1}/(k-h+1)!\,, \tag{3.2}$$

for $k = h-1, h, h+1, \ldots$, respectively.

(ii) Under the condition given by $X(n) = h$ $(h \geq 0)$, $X(n+1)$ can take the values: $X(n+1) = h+1, h, \ldots, 1, 0, -1, \ldots$ with the following probabilities:

① for $1 \leq k \leq h+1$,

$$P_{h,k} = P_{h,h+1-(h-k+1)} = e^{-b} b^{h-k+1}/(h-k+1)!\,, \tag{3.3}$$

② for $k \geq 0$,

$$P_{h,-k} = \int_0^1 \{be^{-bx}(bx)^h/h!\}\{e^{-a(1-x)}(a(1-x))^k/k!\}dx\,. \tag{3.4}$$

3.2 *Invariant distribution*

Let us denote the invariant (stationary) distribution of the skeleton Markov chain $\{X(n)\}_{n=0}^{\infty}$ by $\{\pi_k\}_{k=-\infty}^{\infty}$:

$$\pi_k = \sum_{h=-\infty}^{\infty} \pi_h P_{h,k}\,.$$

(i) For $k > 0$, $P_{h,k}$ vanishes at $h$ if $h \leq k-2$. Therefore, from (3.3), we have

$$\pi_k = \sum_{h=k-1}^{\infty} \pi_h e^{-b} b^{h-k+1}/(h-k+1)! = \sum_{h=0}^{\infty} \pi_{h+k-1} e^{-b} b^h/h!\,. \tag{3.5}$$

Now, suppose

$$\pi_k = cp^k \quad \text{for} \quad k \geq 0\,, \tag{3.6}$$

then, it follows that

$$cp^k = \sum_{h=0}^{\infty} cp^{h+k-1} e^{-b} b^h/h! = cp^{k-1} e^{-b} e^{bp}\,.$$

This leads to

$$p = e^{b(p-1)}, \quad \therefore \quad \log p = b(p-1)\,, \tag{3.7}$$

which has the only solution $p$ $(0 < p < 1)$ for $b > 1$. The value of $c$ will be determined later.

(ii) For $-k \leq 0$, $P_{h,-k}$ vanishes at $h$ if $h \leq -k-2$. Thus, from (3.2) and (3.4), we have

$$(3.8) \qquad \pi_{-k} = \sum_{h=-k-1}^{\infty} \pi_h P_{h,-k}$$
$$= c\int_0^1 be^{b(p-1)x}\{e^{-a(1-x)}(a(1-x))^k/k!\}dx$$
$$+ \sum_{h=0}^{k+1} \pi_{-(k+1-h)} e^{-a}a^h/h! - ce^{-a}a^{k+1}/(k+1)!\,.$$

Let $P(t)$ be the power series with the coefficients $\{\pi_{-k}\}_{k=0}^{\infty}$. Then,

$$(3.9) \qquad P(t) = \sum_{k=0}^{\infty} \pi_{-k}t^k$$
$$= c\int_0^1 be^{b(p-1)x}e^{a(1-x)(t-1)}dx$$
$$+ \{P(t)e^{a(t-1)} - ce^{-a}\}/t - c\{e^{a(t-1)} - e^{-a}\}/t$$
$$= cbe^{a(t-1)}\{e^{b(p-1)-a(t-1)} - 1\}/\{b(p-1) - a(t-1)\}$$
$$+ e^{a(t-1)}\{P(t) - c\}/t\,.$$

This leads to

$$(3.10) \qquad P(t) = c\,\frac{tbp - tbe^{a(t-1)} - \{b(p-1) - a(t-1)\}e^{a(t-1)}}{\{b(p-1) - a(t-1)\}\{t - e^{a(t-1)}\}}\,.$$

By L'Hôpital's theorem, we have

$$(3.11) \qquad P(1) = c\{b(p-1) + a(1-bp)\}/\{b(p-1)(1-a)\}\,.$$

On the other hand, from (3.6), we have

$$(3.12) \qquad P(1) = 1 - \sum_{h=1}^{\infty} \pi_h = 1 - cp/(1-p)\,.$$

Thus, from (3.11) and (3.12), we obtain

$$(3.13) \qquad c = b(1-p)(1-a)/(b-a)\,.$$

Consequently, we can conclude the calculation of the invariant distribution of the skeleton Markov chain $\{X(n)\}_{n=0}^{\infty}$.

## 4. Asymptotic normality of the maximum likelihood estimator

Let $\delta(x) = 1$ if $x > 0$ and $= 0$ if $x \leq 0$. By the ergodicity (1.2) of the stress release process $X(t) = t - N(t)$ and (2.3), we have

$$(4.1)\qquad L(B_T)/T = T^{-1}\int_0^T \delta(X(t-))dt$$

$$\to L_B = \sum_{j=-\infty}^{\infty} \pi_j E\left[\int_0^1 \delta(X(t-))dt \,|\, X(0)=j\right] \qquad \text{(say)}$$

in probability, as $T \to \infty$. Under the condition given by $X(0)=j$, $X(t)$ with $0 \le t \le 1$ can take only the values $X(t) = j+t, j+t-1, j+t-2, \ldots$ and thus, if $j<0$, $X(t)$ is 0 or negative, which implies $\delta(X(t))=0$. Therefore, from (3.6), we see

$$(4.2)\qquad L_B = \sum_{j=0}^{\infty} \pi_j E\left[\int_0^1 \delta(X(t))dt \,|\, X(0)=j\right]$$

$$= c\sum_{j=0}^{\infty} p^j \int_0^1 E[\delta(X(t))\,|\,X(0)=j]dt\,.$$

Now, for $j \ge 0$, it follows from (3.3) that

$$(4.3)\qquad E[\delta(X(t))\,|\,X(0)=j] = \sum_{k=0}^{j} P[X(t)=j+t-k\,|\,X(0)=j]$$

$$= \sum_{k=0}^{j} e^{-bt}(bt)^k/k! = \int_{bt}^{\infty} x^j e^{-x}/j!dx\,.$$

Therefore, we conclude from (4.2) and (4.3) that

$$(4.4)\qquad L_B = c\int_0^1 \left\{\int_{bt}^{\infty} \sum_{j=0}^{\infty} p^j x^j e^{-x}/j!dx\right\} dt = c\int_0^1 \left\{\int_{bt}^{\infty} e^{(p-1)x}dx\right\} dt$$

$$= c/\{b(1-p)\} = (1-a)/(b-a)\,.$$

Since $L(A_T) + L(B_T) = T$, we see

$$(4.5)\qquad L(A_T)/T \to L_A = 1-(1-a)/(b-a) = (b-1)/(b-a) \qquad \text{(say)}\,,$$

in probability, as $T \to \infty$. These, together with (2.6), imply that

$$(4.6)\qquad I_T(\theta)/T \to I(\theta) = \begin{pmatrix} (b-1)/\{a(b-a)\} & 0 \\ 0 & (1-a)/\{b(b-a)\} \end{pmatrix}.$$

We apply the following central limit theorem for the general Poisson-type point process to the proof of the asymptotic normality of the maximum likelihood estimator.

THEOREM 4.1. (*Kutoyants* (1984)) *Let $N(t)$, $0 \le t \le T$, be the point process with the conditional intensity $\lambda_T(t)$, $0 \le t \le T$ and set $M(t) =$*

$N(t)-\int_0^t \lambda_T(s)ds$. *Let* $f_T(t)$, $0 \le t \le T$, *be a predictable function. Suppose the following conditions hold*:

(AN1) $\lim_{T\to\infty}\int_0^T f_T(t)^2\lambda_T(t)dt = \sigma^2 < \infty$, *in probability*.

(AN2) *For any* $\varepsilon > 0$,

$$\lim_{T\to\infty}\int_0^T E\{f_T(t)^2\chi_{\{|f_T(t)|>\varepsilon\}}(t)\lambda_T(t)\}dt = 0 .$$

*Then, it holds that*

$$\mathscr{L}\left[\int_0^T f_T(t)dM(t)\right] \to N(0,\sigma^2), \quad \text{as} \quad T\to\infty . \tag{4.7}$$

Let us denote

$$\begin{aligned}\varDelta_T &= T^{-1/2}(\partial/\partial\theta)\log L_T(\theta)\\ &= \begin{pmatrix} T^{-1/2}\int_0^T \chi_{A_t}(t-)dM(t)/a \\ T^{-1/2}\int_0^T \chi_{B_t}(t-)dM(t)/b\end{pmatrix}.\end{aligned} \tag{4.8}$$

Put $f_T(t) = T^{-1/2}\chi_{A_t}(t)/a$. Then, it holds that (AN1): $T^{-1}\int_0^T \chi_{A_t}(t)/a dt \to (b-1)/\{a(b-a)\}$ in probability, as $T\to\infty$. Since $\{|f_T(t)|>\varepsilon\} = \{|\chi_{A_t}(t)| > \varepsilon a T^{1/2}\}$, (AN2) holds. In the same way, the conditions (AN1) and (AN2) for $\chi_{B_t}(t)$ are checked up. The same discussion holds for the 2-dimensional random vector $\varDelta_T$. Hence, by the above theorem and (4.6), we have

$$\mathscr{L}[\varDelta_T] \to N(0, I(\theta)), \quad \text{as} \quad T\to\infty . \tag{4.9}$$

At the same time, from (2.7) and (4.6), we have the asymptotic normality of the maximum likelihood estimator:

$$\mathscr{L}[T^{1/2}(\hat{\theta}_T-\theta)] \to N(0, I^{-1}(\theta)), \quad \text{as} \quad T\to\infty . \tag{4.10}$$

Similarly, the local asymptotic normality (LAN) of this point process is shown. We see from (2.2) that the log-likelihood ratio for $h=(h_1,h_2)'$ is as follows:

$$\begin{aligned}\Lambda_T(h) &= \log\{L_T(\theta + hT^{-1/2})/L_T(\theta)\}\\ &= \log\{(a+h_1T^{-1/2})/a\}N(A_T) + \log\{(b+h_2T^{-1/2})/b\}N(B_T)\\ &\quad - h_1T^{-1/2}L(A_T) - h_2T^{-1/2}L(B_T)\end{aligned} \tag{4.11}$$

and thus, from (4.6) and by the Taylor's expansion:

$$\log\{(a + h_1 T^{-1/2})/a\} = (h_1/a)T^{-1/2} - (h_1/a)^2 T^{-1} + o(T^{-1})$$

and so on, we have that

$$\Lambda_T(h) = h'\Delta_T(\theta) - G_T(h,\theta) + o_P(1)\,, \tag{4.12}$$

where

$$\begin{aligned} G_T(h,\theta) &= (1/2)h'\begin{pmatrix} T^{-1}L(A_T)/a & 0 \\ 0 & T^{-1}L(B_T)/b \end{pmatrix} h \\ &\to h'I(\theta)h/2 \end{aligned} \tag{4.13}$$

in probability, as $T \to \infty$. This demonstrates that the SSCP process is LAN.

## Acknowledgements

The authors wish to express their gratitude to the referees for their useful comments and kind suggestions.

## References

Hayashi, T. (1986). Laws of large numbers in self-correcting point processes, *Stochastic Process Appl.*, **23**, 319–326.

Hayashi, T. (1988). Local asymptotic normality in self-correcting point processes, *Statistical Theory and Data Analysis* II, (ed. K. Matusita), 551–558, North-Holland, Amsterdam.

Isham, V. and Westcott, M. (1979). A self-correcting point process, *Stochastic Process Appl.*, **8**, 335–347.

Kutoyants, Yu. A. (1984). *Parameter Estimation for Stochastic Process*, (translated by Prakasa Rao, B. L. S.), Heldermann Verlag, Berlin.

Ogata, Y. and Vere-Jones, D. (1984). Inference for earthquake models: a self-correcting model, *Stochastic Process Appl.*, **17**, 337–347.

Snyder, D. L. (1975). *Random Point Processes*, Wiley, New York.

Vere-Jones, D. (1988). On the variance properties of stress release models, *Austral. J. Statist.*, **30A**, 123–135.

Vere-Jones, D. and Ogata Y. (1984). On the moments of a self-correcting process, *J. Appl. Probab.*, **21**, 335–342.