

定常点過程の統計的漸近理論

統計数理研究所 尾形良彦

§1. はじめに

点列 $\omega = \{t_j; j=0, \pm 1, \pm 2, \dots\}$ は集積点をもたない実直線上の点列で $\dots < t_{-1} < 0 \leq t_0 < t_1 < \dots$ となっているものとする。counting measure $N(A) = N(A, \omega)$ はボレル集合 $A \in \mathcal{B}^1$ に対して $\omega \cap A$ の点の数を対応させるものである。点過程の法則 P が与えられたとき時間 t についての complete intensity function は次のように定義される。

$$(1.1) \quad \lambda(t, \omega) = \lim_{\delta \rightarrow 0} \frac{1}{\delta} P\{N(t, t+\delta)=1 \mid \mathcal{H}_t\}$$

ここに $\mathcal{H}_t$ は ω の時間 t 以前の履歴、すなわち $N(A)$, $A \in \mathcal{B}_t$ で生成された σ -加法族である。いかえると $\lambda(t, \omega)$ は t と $\{t_j \in \omega, t_j < t\}$ を変数とする関数として表現される。

以下いくつかの例をあげる。

例1 ポアソン過程

$$\lambda(t, \omega) = \lambda(t)$$

このとき $N(A)$ の分布は, $k=0, 1, 2, \dots$ に対して

$$P(N(A)=k) = (1/k!) \left(\int_A \lambda(t) dt \right)^k \exp - \int_A \lambda(t) dt$$

定常ならば $\lambda(t) = \lambda = \text{const.}$

例2 Survivor function $F(x)$ の更新過程

$$\lambda(t, \omega) = - \frac{d}{dt} \log F(t - t^*)$$

ここで $t^* < t$ は最近の点。 (hazard function)

例3 Wold process [6]

これは更新過程の拡張になっている。かんたんのため $m=2$ の場合を与える。時間 t について最近の2点 $t^{**} < t^* < t$ についての区間の長さに対する同次分布を $F(t-t^*, t^*-t^{**})$ で与えると、この条件付 hazard 関数は

$$\lambda(t, \omega) = - \frac{\partial}{\partial t} \left\{ \log \frac{\partial}{\partial \sigma} \left(\frac{\partial}{\partial \sigma} - \frac{\partial}{\partial t} \right) F(\tau, \sigma) \right\}$$

となる。

例4 Hawkes の self-exciting process [2]

出生死滅過程の死滅がない場合について移住時点と出生時点がつくる点列を考える。移住者が到着するのは平均 μ の定常ポアソンに従い、それらおよびそれらの子孫たちは生まれてからの時間 $(t, t+dt)$ のあいだに $\gamma(t) dt$ の確率で子どもを一つ生む。 $\int_0^\infty \gamma(u) du < 1$ であるとき点過

程は定常になり、intensity function は

$$\lambda(t, \omega) = \mu + \int_{-\infty}^t \gamma(t-u) dN(u) = \mu + \sum_{t_i < t} \gamma(t-t_i)$$

で与えられる。

いま、与えているクラスの点過程は、complete intensity function に対して^{定常}点過程が存在し、1かき一意にきまるものに限定しておく。上にあげた例たちはこのことを保証されている。さて時間区間 $[0, T]$ に於いて点列 $\{t_i\}_{t \geq 0}$ が観測されたときに我々は次のような尤度関数を考えることができる。すなわち intensity function をパラメライズして $\lambda_\theta(t, \omega)$, $\theta \in \Theta$, としたとき、対数尤度関数は形式的に次で与えられる。

$$\begin{aligned} \mathcal{L}_T(\theta) &= - \int_0^T \lambda_\theta^*(t, \omega) dt + \int_0^T \log \lambda_\theta^*(t, \omega) dN(t) \\ &= - \int_0^T \lambda_\theta^*(t, \omega) dt + \sum_{0 \leq t_i \leq T} \log \lambda_\theta^*(t_i, \omega) \end{aligned}$$

ただし、実際的には

$$\lambda_\theta^*(t, \omega) = E\{\lambda_\theta(t, \omega) | \mathcal{H}_{0,t}\}$$

となるために λ_θ^* は定常でないが、 $T \rightarrow \infty$ とともに漸近的には λ_θ のふるまいと遠ゆりのことが期待される。このことには最後の節でのべる。ここでは λ_θ^* は λ_θ と思って話をすすめる。

最尤推定量 $\hat{\theta}_T = \hat{\theta}_T(t_i; 0 \leq t_i \leq T)$ は尤度を最大にするものとして定義する。 $L_T(\hat{\theta}_T) = \max_{\theta} L_T(\theta)$ 。
 これはまた $\partial L_T(\theta)/\partial \theta = 0$ の解にもなっている。

§2. 仮定および準備。

仮定

- (A1) 点過程は定常エルゴード的。
- (A2) 点過程は orderly; $\lim_{\delta \rightarrow 0} \frac{1}{\delta} P\{N(0, \delta) \geq 2\} = 0$ 。
- (A3) $\sup_{0 < \delta \leq 1} \frac{1}{\delta} E\{N(0, \delta)^2 | \mathcal{H}_{-\infty, 0}\}$ は P -可積分。
- (A4) Θ は compact, $\Theta \subset \mathbb{R}^d$
- (A5) 任意の $\theta \in \Theta, t \in \mathbb{R}^+$ に対して $\lambda_{\theta}(t, \omega) > 0$ a.s.
- (A6) $\theta_1 = \theta_2$ if and only if $\lambda_{\theta_1}(0, \omega) = \lambda_{\theta_2}(0, \omega)$ a.s. P 。
- (A7) $\forall \theta \in \Theta, t \in \mathbb{R}$ に対して $\partial \log \lambda / \partial \theta_i, \partial^2 \log \lambda / \partial \theta_i \partial \theta_j, \partial^3 \log \lambda / \partial \theta_i \partial \theta_j \partial \theta_k, i, j, k = 1, 2, \dots, d$, が存在する。
- (A8) 任意 $t \in \mathbb{R}, \theta \in \Theta$ について $|\partial \lambda / \partial \theta_i| \leq \Lambda_1(t, \omega), |\partial \lambda / \partial \theta_i \partial \theta_j| \leq \Lambda_2(t, \omega), i, j = 1, 2, \dots, d$, で $\Lambda_1(t, \omega), \Lambda_2(t, \omega)$ は $\mathcal{H}_{-\infty, t}$ -adapted かつ P に依りて 2 乗可積分。
- (A9) 任意の θ に対して $I(\theta) = \{E(\frac{1}{\lambda} \frac{\partial \lambda}{\partial \theta_i} \frac{\partial \lambda}{\partial \theta_j})\}_{i, j=1, \dots, d}$ は正則行列で $\frac{1}{\lambda} \frac{\partial \lambda}{\partial \theta_i} \frac{\partial \lambda}{\partial \theta_j}, i, j = 1, \dots, d$, は P に依りて 2 乗可積分。
- (A10) 任意の $\theta \in \Theta, t \in \mathbb{R}$ に対して $|\partial^3 \lambda / \partial \theta_i \partial \theta_j \partial \theta_k| \leq$

×

$H(t, \omega)$, $|\partial^3 \log \lambda / \partial \theta_i \partial \theta_j \partial \theta_k| \equiv G(t, \omega)$, $i, j, k = 1, \dots, d$,
 であり適当な正数 $M > 0$ に對して $E_0 \{ H(t, \omega) \} < M$,
 $E_0 \{ \lambda_0(t, \omega)^3 G(t, \omega)^2 \} < M$.

補題 1

(A1) ~ (A3) のもとで

$$1^\circ. E[N(0, 1)^2] < \infty$$

$$2^\circ. \lim_{\delta \rightarrow 0} \frac{1}{\delta} P\{N(t, t+\delta) \geq 2 \mid \mathcal{H}_t\} = 0$$

$$3^\circ. \lim_{\delta \rightarrow 0} \frac{1}{\delta} E[N(t, t+\delta)^2 \mid \mathcal{H}_t] = \lim_{\delta \rightarrow 0} \frac{1}{\delta} E[N(t, t+\delta) \mid \mathcal{H}_t]$$

証明

1°. $\delta = 1$ とし (A3) から $t, t+\delta$ に $\mathcal{G}_t$ かかる。

$$2^\circ. E\left[\frac{1}{\delta} \cdot P\{N(t, t+\delta) \geq 2 \mid \mathcal{H}_t\}\right]$$

$$\leq E\left[\frac{1}{\delta} \sum_{i=1}^{\infty} i^2 P\{N(t, t+\delta) = i \mid \mathcal{H}_t\}\right] \leq E\left[\frac{1}{\delta} E\{N(t, t+\delta) \mid \mathcal{H}_t\}\right]$$

だから (A3) による orderliness (A2) とあわせて

$$E\left[\lim_{\delta \rightarrow 0} \frac{1}{\delta} P\{N(t, t+\delta) \geq 2 \mid \mathcal{H}_t\}\right]$$

$$= \lim_{\delta \rightarrow 0} \frac{1}{\delta} P\{N(t, t+\delta) \geq 2\} = 0$$

被積分関数はいたるところ non-negative なるから

$$\lim_{\delta \rightarrow 0} P\{N(t, t+\delta) \geq 2 \mid \mathcal{H}_t\} = 0.$$

3° orderliness (A3) に對して $n \rightarrow \infty$ に對して

$$\sum_{k=0}^{n-1} N\left(\frac{k}{n}, \frac{k+1}{n}\right)^2 - \sum_{k=0}^{n-1} N\left(\frac{k}{n}, \frac{k+1}{n}\right) \rightarrow 0 \quad \text{a.s.}$$

$$N(0,1) = \sum_{k=0}^{n-1} N\left(\frac{k}{n}, \frac{k+1}{n}\right) \quad \text{である}$$

$$\sum_{k=0}^{n-1} N\left(\frac{k}{n}, \frac{k+1}{n}\right)^2 \leq \left\{ \sum_{k=0}^{n-1} N\left(\frac{k}{n}, \frac{k+1}{n}\right) \right\}^2 = N(0,1)^2$$

1° に對して有界収束定理から

$$n E\left[N\left(0, \frac{1}{n}\right)^2\right] \rightarrow E\{N(0,1)\}$$

$$n E\left[N\left(0, \frac{1}{n}\right)\right] \rightarrow E\{N(0,1)\}.$$

すなわち

$$\lim_{\delta \rightarrow 0} \frac{1}{\delta} E\{N(t, t+\delta)^2\} = \lim_{\delta \rightarrow 0} \frac{1}{\delta} E\{N(t, t+\delta)\} = E[N(t,1)]$$

$N(t, t+\delta)$ は ^{非負} integer-valued である。3 次の非積分関数は

は non-negative である。

$$E\left[\lim_{\delta \rightarrow 0} \frac{1}{\delta} E\{N(t, t+\delta)^2 | \mathcal{H}_t\} - \lim_{\delta \rightarrow 0} \frac{1}{\delta} E\{N(t, t+\delta) | \mathcal{H}_t\}\right]$$

$$= \lim_{\delta \rightarrow 0} \left[\frac{1}{\delta} E\{N(t, t+\delta)^2\} - \frac{1}{\delta} E[N(t, t+\delta)] \right] = 0$$

だから 3° が導かれる。

§3. 結果

我々は確率過程 $\xi = \{\xi(t, \omega), t \geq 0\}$ が adapted であるとは、固定した $t \geq 0$ に對して $\xi(t, \omega)$ が $\mathcal{H}_{\infty, t^-}$ 可測

であるときに言う。ここでは定義を与えなから *adapted* な
 確率過程の subclass Φ には predictable な確率過程
 (c.f. [3] page 2) のことを意味する。ここでは、我々は
 $\xi = \{\xi(t, \omega); t \geq 0\}$ から a.s. な ω に対して標本関数が左連
 続ならば $\Phi \ni \xi$ である事実も知るために満足する = とし
 よう。仮定 α のときに述べるのを忘れてしまつたことであるから
 $\lambda_0(t, \omega), \frac{\partial}{\partial \theta_i} \lambda_0(t, \omega), \frac{\partial^2}{\partial \theta_i \partial \theta_j} \lambda_0(t, \omega)$ などの intensity
 process の必要な汎関数はすべて predictable であること
 とする。

さて、これを我々はすでに書いてしまつたことであるから
 確率積分 (Stieltjes) $\int_0^T \xi(t, \omega) dN(t) = \sum_{0 \leq t_i \leq T} \xi(t_i(\omega), \omega)$ は、
 ここでは predictable な ξ に対して考へてゐる。そうすると
 我々は次のような形式的な演算をすることを許さへてゐる。

$$\int_0^T E\{\lambda_0(t, \omega) | \xi(t, \omega)\} dt < \infty \text{ ならば } \\
 E\left\{\int_0^T \xi(t, \omega) dN(t)\right\} = E\left[E\left\{\int_0^t \xi(t, \omega) dN(t) | \mathcal{H}_{-\infty, t}\right\}\right] = E\left[\int_0^t \xi(t, \omega) E\{dN(t) | \mathcal{H}_{-\infty, t}\}\right] = E\left[\int_0^T \xi(t, \omega) \lambda_0(t, \omega) dt\right]$$

以下、結果をあげて証明のあらすじを述べる。

定理 1

仮定 (A1) から (A9) のもとに: $\theta = \theta_0$ に対して

$$\begin{aligned}
 E\left\{\frac{\partial \mathcal{L}_T}{\partial \theta_i}\right\} &= 0, \quad i=1, 2, \dots, d \\
 E\left\{\frac{\partial \mathcal{L}_T}{\partial \theta_i} \frac{\partial \mathcal{L}_T}{\partial \theta_j}\right\} &= -E\left\{\frac{\partial^2 \mathcal{L}_T}{\partial \theta_i \partial \theta_j}\right\} = T \cdot E\left\{\frac{1}{\lambda} \frac{\partial \lambda}{\partial \theta_i} \frac{\partial \lambda}{\partial \theta_j}\right\} \quad i, j=1, \dots, d.
 \end{aligned}$$

証明.

後生を示す。 $\theta = \theta_0$ に於て補題 1 の 1°, 2° に従って

$$\begin{aligned} E \left[\frac{\partial^2 \mathcal{L}_T}{\partial \theta_i \partial \theta_j} \right] &= E \left[- \int_0^T \frac{\partial^2 \lambda}{\partial \theta_i \partial \theta_j} dt + \int_0^T \frac{1}{\lambda} \frac{\partial^2 \lambda}{\partial \theta_i \partial \theta_j} dN(t) - \int_0^T \frac{1}{\lambda^2} \frac{\partial \lambda}{\partial \theta_i} \frac{\partial \lambda}{\partial \theta_j} dN(t) \right] \\ &= E \left\{ - \int_0^T \frac{1}{\lambda} \frac{\partial \lambda}{\partial \theta_i} \frac{\partial \lambda}{\partial \theta_j} dt \right\} = -T \cdot E \left\{ \frac{1}{\lambda} \frac{\partial \lambda}{\partial \theta_i} \frac{\partial \lambda}{\partial \theta_j} \right\}. \end{aligned}$$

他方, $\theta = \theta_0$ に於て

$$\begin{aligned} &E \left\{ \frac{\partial \mathcal{L}_T}{\partial \theta_i} \frac{\partial \mathcal{L}_T}{\partial \theta_j} \right\} \\ &= E \left[\int_0^T \int_0^T \frac{\partial \lambda(s)}{\partial \theta_i} \frac{\partial \lambda(t)}{\partial \theta_j} \left\{ ds dt - \frac{dN(s)dt}{\lambda(s)} - \frac{ds dN(t)}{\lambda(t)} + \frac{dN(s)dN(t)}{\lambda(s)\lambda(t)} \right\} \right] \\ &= E \left[\iint_{\{0 \leq s < t \leq T\}} + \iint_{\{0 \leq t < s \leq T\}} + \iint_{\{s=t\}} \right] \\ &= I_1 + I_2 + I_3, \end{aligned}$$

こゝに $\lambda(s) = \lambda_{\theta_0}(s, \omega)$ 。 $s < t$ に於て π_t 式で表すにいと

$E \{ dN(s) dN(t) | \mathcal{H}_t \} = dN(s) \lambda(t) dt$ とあるから $I_1 = 0$

同様にして $I_2 = 0$ 。最後に補題 1 の 3° に従って π_t 式で表すにいと

$$E \{ (dN(t))^2 | \mathcal{H}_t \} = E \{ dN(t) | \mathcal{H}_t \} = \lambda(t) dt$$

から成るから

$$I_3 = E \left\{ \int_0^T \frac{1}{\lambda} \frac{\partial \lambda}{\partial \theta_i} \frac{\partial \lambda}{\partial \theta_j} dt \right\} = T \cdot E \left[\frac{1}{\lambda} \frac{\partial \lambda}{\partial \theta_i} \frac{\partial \lambda}{\partial \theta_j} \right].$$

以上により被積分関数が non-anticipating であるから

と仮定したから明らかである。

補題2

確率過程 $\xi(t, \omega)$ は定常で2次元 $- \times 2$ ト $\mathbb{R}^4$ 上 non-anticipating であるとする。すると

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \xi(t, \omega) dt = E \{ \xi(0, \omega) \} \quad \text{a.s.},$$

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \xi(t, \omega) \frac{dN(t)}{\lambda(t)} = E \{ \xi(0, \omega) \} \quad \text{a.s.}.$$

証明

補題の仮定と (A1) によって $\xi(t, \omega)$ はエルゴード的であるから最初のものはエルゴード定理。2 番目について、

$$\eta(T, \omega) = \int_0^T \xi(t, \omega) \frac{dN(t)}{\lambda(t)} - \int_0^T \xi(t, \omega) dt$$

とあけ。

$$\gamma_i = \eta(i, \omega) - \eta(i-1, \omega), \quad i=1, 2, \dots, [T]$$

は定常なマルコフチェーンの差になっている。よって Kolmogorov の不等式によつて $\sum_{i=1}^{[T]} \gamma_i / T \rightarrow 0 \quad \text{a.s.}$ となる。故に $\eta(T, \omega) / T \rightarrow 0 \quad \text{a.s.}$ となり、前半の結果とあわせて後半が得られた。

補題3

単位区間 $[0, 1]$ における対数尤度比

$$L_1^*(\theta) = \int_0^1 (\lambda_\theta - \lambda_{\theta_0}) dt + \int_0^1 \log \frac{\lambda_{\theta_0}}{\lambda_\theta} dN(t)$$

に対し

$$E [L_1^*(\theta)] \geq 0, \quad \theta \in \Theta$$

したがって等号は $\lambda_\theta(0, \omega) = \lambda_{\theta_0}(0, \omega)$ a.s. のときにかぎる。

証明

$$E[L_1^*(\theta)] = E\left[\lambda_\theta(0, \omega) - \lambda_{\theta_0}(0, \omega) + \lambda_{\theta_0}(0, \omega) \log \frac{\lambda_{\theta_0}(0, \omega)}{\lambda_\theta(0, \omega)}\right]$$

一般に $x > 0$ に対して $\log x \geq 1 - x^{-1}$ が成立し
て等号は $x = 1$ のときにかぎることにより証明できる。

定理 2.

仮定 (A1) ~ (A7) のもとに最尤推定量 $\hat{\theta}_T = \hat{\theta}_T(x_i, 0 \leq x_i \leq T)$ は一貫性をもつ。

証明

θ の近傍 U から $\{\theta\}$ に縮まると同時に、仮定 (A7) から λ_θ は θ について連続であるから、

$$E\left[\inf_{\theta' \in U} \lambda_{\theta'}\right] \rightarrow E[\lambda_\theta]$$

$$E\left[\lambda_{\theta_0} \log \frac{\lambda_{\theta_0}}{\sup_{\theta' \in U} \lambda_{\theta'}}\right] \rightarrow E\left[\lambda_{\theta_0} \log \frac{\lambda_{\theta_0}}{\lambda_\theta}\right]$$

となる。いず θ_0 の近傍を任意に U_0 ととる。任意の $\theta \in \Theta \setminus U_0$ に対して $E[L_1^*(\theta)] \geq 3\varepsilon$ となるような適当な $\varepsilon > 0$ をとることができる。 $L_1^*(\theta)$ は θ について連続だから補題 3 と仮定 (A6) をつかえばよいからである。

さて任意の $\theta \in \Theta \setminus U_0$ に対して θ の近傍 U を適当にとれば、

$$E\left[\inf_{\theta' \in U} \lambda_{\theta'} - \lambda_{\theta_0} + \lambda_{\theta_0} \log \frac{\lambda_{\theta_0}}{\sup_{\theta' \in U} \lambda_{\theta'}}\right] \geq L_1^*(\theta) - \varepsilon \geq 2\varepsilon$$

とできる。④ \setminus U_0 はコンパクト集合であるからこれを覆う有限個の Θ_s , $s=1,2,\dots,N$ を選ぶことができる。すなわち $U_s = U(\Theta_s)$ とおけば $\textcircled{4} \setminus U_0 \subset \bigcup_{s=1}^N U_s$ 。ところで $\inf_{\Theta \in U} \lambda_{\Theta}(t, \omega)$, $\sup_{\Theta \in U} \lambda_{\Theta}(t, \omega)$ は predictable process だから T を充分大きくとる λ 補足 2 に従って $\varepsilon > 0$ あり $s=1,2,\dots,N$ に対して $(T \geq T_0(\varepsilon))$

$$\begin{aligned} & \frac{1}{T} L_T(\Theta_0) - \sup_{\Theta \in U_s} \frac{1}{T} L_T(\Theta) \\ & \geq \frac{1}{T} \int_0^T (\inf_{\Theta \in U_s} \lambda_{\Theta} - \lambda_{\Theta_0}) dt + \frac{1}{T} \int_0^T \log \frac{\lambda_{\Theta_0}}{\sup_{\Theta \in U_s} \lambda_{\Theta}} dN(t) \\ & \geq E \left[\inf_{\Theta \in U_s} \lambda_{\Theta} - \lambda_{\Theta_0} + \lambda_{\Theta_0} \log \frac{\lambda_{\Theta_0}}{\sup_{\Theta \in U_s} \lambda_{\Theta}} \right] + \varepsilon \\ & \geq \varepsilon \end{aligned}$$

が成立する。すなわち Θ_0 の任意の近傍 U_0 について $T_0 = T_0(U_0)$ が存在して、 $T \geq T_0$ ならば

$$\sup_{\Theta \in U_0} L_T(\Theta) \geq \sup_{\Theta \in \Theta - U_0} L_T(\Theta) + \varepsilon T \quad \text{a.s.}$$

これは定理 2 を導く。

定理 3

Θ_0 の適当な近傍で尤度の Hessian $\left\{ \frac{1}{T} \frac{\partial^2 L_T(\Theta)}{\partial \theta_i \partial \theta_j} \right\}_{i,j=1,2,\dots,d}$ は漸近的に negative-definite.

証明略。

例 5. intensity function の parametrization と尤度の単峰性。

$$\lambda_\theta(t, \omega) = \sum_{i=1}^k \theta_i \xi_i(t, \omega) + \eta(t, \omega)$$

$$\lambda_\theta(t, \omega) = \exp \left\{ \sum_{i=1}^k \theta_i \xi_i(t, \omega) + \eta(t, \omega) \right\}$$

(ただし ξ_i, η は点過程 ω について predictable)

は尤度は単峰である。事実, それぞれ $u_i \in \mathbb{R} \quad i=1, 2, \dots, d$

に対して

$$\sum_{i,j=1}^d u_i u_j \frac{\partial^2 \mathcal{L}_T}{\partial \theta_i \partial \theta_j} = \begin{cases} - \int_0^T \left\{ \left(\sum_{i=1}^k u_i \xi_i \right)^2 / \lambda_\theta^2 \right\} dN(t) \\ - \int_0^T \left(\sum_{i=1}^k u_i \xi_i \right)^2 \lambda_\theta dt \end{cases}$$

となる。

定理 4

$$\frac{1}{\sqrt{T}} \frac{\partial \mathcal{L}_T(\theta_0)}{\partial \theta} \rightarrow N(0, I(\theta_0)), \quad T \rightarrow \infty$$

証明。

$0 \leq S \leq T, i=1, 2, \dots, d$ について

$$E \left\{ \frac{\partial \mathcal{L}_T(\theta_0)}{\partial \theta_i} \middle| \mathcal{H}_S \right\} = \frac{\partial \mathcal{L}_S(\theta_0)}{\partial \theta_i} + E [\Delta(S, T) | \mathcal{H}_S]$$

ただし

$$E \{ \Delta(S, T) | \mathcal{H}_S \} = E \left[\int_S^T \frac{\partial \lambda}{\partial \theta} \left\{ \frac{dN(t)}{\lambda} - dt \right\} \middle| \mathcal{H}_S \right] = 0$$

となる。よって $\partial \mathcal{L}_T(\theta_0) / \partial \theta$ は $\mathcal{H}_T$ に関するマルコフ過程である。

□

$$\frac{\partial \mathcal{L}_T(\theta_0)}{\partial \theta} = \sum_{k=1}^{[T]} \Delta(k-1, k) + \Delta([T], T)$$

と分解すると確率変数列 $\{\Delta(k-1, k)\}_{k=1,2,\dots}$ は定常エルゴード的なマルコフ連鎖に差になり $E\{\Delta(0,1)\Delta(0,1)'\} = I(\theta_0)$ とするこゝから定理1より導かれるから

$$\frac{1}{\sqrt{[T]}} \sum_{k=0}^{[T]-1} \Delta(k, k+1) \longrightarrow N(0, I(\theta_0)).$$

一方 $1/\sqrt{T} \cdot \Delta([T], T) \rightarrow 0$ in prob. から定理4を得る。

定理5

$\hat{\theta}_T$ が最尤推定量のとき $T \rightarrow \infty$ とともに

$$\sqrt{T}(\hat{\theta}_T - \theta_0) \longrightarrow N(0, I(\theta_0)^{-1})$$

$$2\{\mathcal{L}_T(\hat{\theta}_T) - \mathcal{L}_T(\theta_0)\} \longrightarrow \chi^2_\alpha$$

証明

$$0 = \frac{1}{\sqrt{T}} \frac{\partial \mathcal{L}_T(\theta_0)}{\partial \theta} + \frac{1}{T} \frac{\partial^2 \mathcal{L}_T(\theta_0)}{\partial \theta \partial \theta'} \sqrt{T}(\hat{\theta}_T - \theta_0) + \sqrt{T}(\hat{\theta}_T - \theta_0)' \left\{ -\frac{\alpha}{T} \int_0^T H(t, \omega) dt + \frac{\beta}{T} \int_0^T G(t, \omega) dN(t) \right\} (\hat{\theta}_T - \theta_0)$$

は仮定(A7)~(A10)と定理2によって充分大きい T に対して導かれる。ここに α, β は確率変数も成る $d \times d$ 行列で $|\alpha_{ij}|, |\beta_{ij}| \leq d^2/2$ 。任意の $\delta > 0$ に対して T を充分大きくとったとき $\theta = \theta_0$ に対して

$$\left| \frac{1}{T} \int_0^T \frac{\partial \lambda}{\partial \theta \partial \theta'} dt + \frac{1}{T} \int_0^T \frac{1}{\lambda} \frac{\partial \lambda}{\partial \theta} \frac{\partial \lambda}{\partial \theta'} dN(t) \right| < \delta d^2$$

$$\left| I(\theta_0) + \frac{1}{T} \int_0^T -\frac{1}{\lambda^2} \frac{\partial \lambda}{\partial \theta} \frac{\partial \lambda}{\partial \theta} dN(t) \right| < \delta d^2$$

$$\left| \frac{\alpha}{T} \int_0^T H(t, \omega) dt + \frac{\beta}{T} \int_0^T G(t, \omega) dN(t) \right| < \varepsilon d^3 M^3$$

となる。 $T \rightarrow \infty$ と ε ともに $|\hat{\theta}_T - \theta_0| \rightarrow 0$ a.s. になる。だから適当

な $\varepsilon_T > 0$ 但し $T \rightarrow \infty$ と ε ともに $\varepsilon_T \rightarrow 0$ に選んで

$$\left| \frac{1}{\sqrt{T}} \frac{\partial \mathcal{L}_T(\theta_0)}{\partial \theta} - \sqrt{T} I(\theta_0) (\hat{\theta}_T - \theta_0) \right| \leq \varepsilon_T |\sqrt{T} (\hat{\theta}_T - \theta_0)|$$

を得るので前半の結果が得える。後半については、

$$2 \{ \mathcal{L}_T(\hat{\theta}_T) - \mathcal{L}_T(\theta_0) \}$$

$$= 2 \frac{\partial \mathcal{L}_T(\theta_0)}{\partial \theta'} (\hat{\theta}_T - \theta_0) + (\hat{\theta}_T - \theta_0)' \frac{\partial^2 \mathcal{L}_T(\theta_0)}{\partial \theta \partial \theta'} (\hat{\theta}_T - \theta_0) +$$

$$+ |\hat{\theta}_T - \theta_0|^3 \left\{ \int_0^T \alpha^* H(t, \omega) dt + \int_0^T \beta^* G(t, \omega) dN(t) \right\}$$

となるように α^*, β^* , $|\alpha^*|, |\beta^*| \leq d^2/2$ を選べし

最後の項は 0 に収束 (a.s.) するので後半の結果を得る。

4. 初期条件についての問題

いままで漸近的な結果を出すために intensity function $\lambda_\theta(t, \omega)$ は定常過程である必要があった。しかしこれは、とりまきおさず $\lambda_\theta(t, \omega)$ が $\mathcal{H}_t = \mathcal{H}_{-\infty, t}$ の関数である

こと、すなわち $-\infty$ の過去からの影響があることを考慮にい
 れているということである。実際には観測できるのは時間
 区間 $[0, T]$ 中のため、当然 intensity function とし
 て、たとえば $\lambda_0^*(t, \omega) = E[\lambda_0(t, \omega) | \mathcal{H}_{0,t}]$ を代用す
 ることになる。これは $T \rightarrow \infty$ とともに $\lambda(t, \omega)$ に収束す
 ることはわかるが、問題は、そのスピードである。その違
 いが、補題2のタイプの極限定理や中心極限定理
 に影響を与えない程度のものである必要がある。
 確率収束のいみでの条件を要求するとき下の条件は見
 容の標準になるものと思う。

条件B

$$|\lambda_0(t, \omega) - \lambda_0^*(t, \omega)| \leq f_0(t, \theta) \xi_0(\theta, \omega),$$

$$\left| \frac{\partial \lambda}{\partial \theta_i} - \frac{\partial \lambda^*}{\partial \theta_i} \right| \leq f_1(t, \theta) \xi_1(\theta, \omega),$$

$$\left| \frac{\partial \lambda}{\partial \theta_i \partial \theta_j} - \frac{\partial \lambda^*}{\partial \theta_i \partial \theta_j} \right| \leq f_2(t, \theta) \xi_2(\theta, \omega)$$

$$\left| \frac{\partial^3 \lambda}{\partial \theta_i \partial \theta_j \partial \theta_k} - \frac{\partial^3 \lambda^*}{\partial \theta_i \partial \theta_j \partial \theta_k} \right| \leq f_3(t, \theta) \xi_3(\theta, \omega)$$

$i, j, k = 1, \dots, d.$

ただし $T \rightarrow \infty$ とともに

$$\frac{1}{\sqrt{T}} \int_0^T f_\nu(t, \theta) dt \rightarrow 0, \quad \nu = 0, 1, 2, 3,$$

これらの条件は、例1ならば定常ポアソン、例2,3は自動的に満足している。例4についてまたとせば、

$$\lambda_0 = \mu + \int_{-\infty}^t \alpha e^{-\beta(t-u)} dN(u), \quad 0 < \alpha < \beta,$$

ならば

$$\lambda_0^* = \mu + \frac{\alpha\mu}{\beta-\alpha} e^{-\beta t} + \int_0^t \alpha e^{-\beta(t-u)} dN(u)$$

だから

$$|\lambda_0 - \lambda_0^*| = e^{-\beta t} \left\{ \frac{\alpha\mu}{\beta-\alpha} + \int_{-\infty}^0 \alpha e^{\beta u} dN(u) \right\}$$

である。そして $f_2(t, 0) = (-x)^{\nu} e^{-\beta t}$ となって条件が満たされていることになる。

§5. Cramer-Rao の不等式

いま区間 $(0, T)$ において点列 $0 \leq t_1 < t_2 < \dots < t_n \leq T$ が観測されたとして点過程 P_0 の母数 $\theta \in \Theta \subset \mathbb{R}^k$ に対する推定量 $\delta_T = \delta_T(\omega)$, $(\omega = (\dots, t_1, t_2, \dots, t_n, \dots))$, を考える。さてここで点過程 P_θ , $\theta \in \Theta$, は区間 $(0, T)$ で intensity 1 のポアソン過程 P_π に対して絶対連続であるとする。このとき Radon-Nykodim 微分は

$$P_\theta(\omega) = \frac{dP_\theta}{dP_\pi} = \exp \left\{ \int_0^T \log \lambda_\theta^*(t) dN(t) + \int_0^T (1 - \lambda_\theta^*(t)) dt \right\}$$

で与えられることが知られている。以下において必要な正則条件（たとえば微分と積分の交換など）は、すべて満す

ものとする。

定理 6

$$E_0(\delta_T) = \theta + b_T(\theta), \quad I_T^*(\theta) = \int_0^T E\left\{\frac{1}{X^*(t)} \frac{\partial \lambda^*}{\partial \theta} \frac{\partial \lambda^*}{\partial \theta'}\right\} dt$$

$$\Sigma(\delta_T) \geq \left(I + \frac{db_T(\theta)}{d\theta}\right)' I_T^*(\theta)^{-1} \left(I + \frac{db_T(\theta)}{d\theta}\right)$$

ここで $\Sigma(\delta_T)$ は δ_T の共分散行列。等号は $\delta_T = \text{const} \times \frac{\partial}{\partial \theta} \log p_0(\omega)$ の時。不等号は非負定値を意味する。

証明

$$E_0(\delta_T) = \int \delta_T P_0(d\omega) = \int \delta_T(\omega) p_0(\omega) P_\pi(d\omega).$$

$$\begin{aligned} \therefore \frac{d}{d\theta} E_0(\delta_T) &= \int \delta_T \frac{\partial}{\partial \theta} p_0(\omega) P_\pi(d\omega) = \int \delta_T \left(\frac{\partial}{\partial \theta} \log p_0(\omega) \right) p_0(\omega) P_\pi(d\omega) \\ &= \int \delta_T \frac{\partial}{\partial \theta} \log p_0(\omega) P_0(d\omega) = E_0 \left[\delta_T \frac{\partial}{\partial \theta} \log p_0(\omega) \right] \end{aligned}$$

定理 1 と同様に $E_0 \left[\frac{\partial}{\partial \theta} \log p_0(\omega) \right] = 0$ かつ ε と s の内積 $s, t \in R^k$ に対して

$$\begin{aligned} \left\{ t' \left(I + \frac{d}{d\theta} b_T(\theta) \right) s \right\}^2 &= \left\{ t' \cos \left(\delta_T \frac{\partial}{\partial \theta} \log p_0(\omega) \right) s \right\}^2 \\ &\leq E \left(t' (\delta_T - \theta - b_T(\theta)) (\delta_T - \theta - b_T(\theta))' t \right) E \left[s' \frac{\partial}{\partial \theta} \log p_0(\omega) \frac{\partial}{\partial \theta} \log p_0(\omega) s \right] \end{aligned}$$

ここで $S = E \left[(\delta_T - \theta - b_T(\theta)) (\delta_T - \theta - b_T(\theta))' \right]^{-1} \left(I + \frac{d}{d\theta} b_T(\theta) \right) t$ と置いて代入すると証明は終る。

さて $I_T^*(\theta)/T \rightarrow I(\theta)$, $T \rightarrow \infty$, となることから

λ^* と λ の条件からわかるから、この定理 6 と定理 5 によって次の定理を得る。

定理7

仮定 A と条件 B のもとに最尤推定量 $\hat{\theta}_T$ は Best Asymptotic Normal 推定量である。

例5

ポアソン過程の intensity θ についての Fisher-情報量は $1/\theta$ である。最尤推定量は $N(0, T)/T$ であり、 $(0, \infty)$ で一様な事前分布を与えた バイス推定量は $(N(0, T)+1)/T$ となる。これは、いずれも Cramer-Rao の lower-bound を attain するから前者が不偏推定量である。

§6. ポアソン過程の最尤推定量による特徴づけ

更新過程も生き残り関数によって以下のようにパラメータ化する。すなわち、いま $\int_0^\infty dF(x) = 1$, $\int_0^\infty x dF(x) = 1$ なる分布関数に対して生き残り関数が $f_0(x) = 1 - F(\theta x)$ なる更新過程を考えたとき $E[N(0, 1)] = 0$ となることが知られている。 $F(x)$ は必要な正則条件をすべてみたすとする。

定理8

更新過程の平均 intensity θ についての最尤推定量が $N(0, T)/T$ であるための必要十分条件は $F(x) = e^{-x}$ である。

定常ポアソン過程となることである。

証明

十分条件は前節の例5で与えた。必要条件。いま観測
 $0 \leq t_1 < t_2 < \dots < t_n \leq T$ を任意に与えたとき 対数尤度は

$$\log L_T(\theta) = \sum_{i=1}^n \log \theta f(\theta(t_i - t_{i-1})) + \log \{1 - F(\theta(T - t_n))\}$$

で与えられることが計算できる。(c.f. 例2) このとき
 $(\partial/\partial \theta) \log L_T(\frac{n}{T}) = 0$ が必要条件から要求されて、これから

$$0 = T + \sum_{i=1}^n (t_i - t_{i-1}) \frac{f'(\frac{n}{T}(t_i - t_{i-1}))}{f(\frac{n}{T}(t_i - t_{i-1}))} + \frac{-(T - t_n) f(\frac{n}{T}(T - t_n))}{1 - F(\frac{n}{T}(T - t_n))}$$

ただし $f(x) = dF(x)/dx$, $t_0 = 0$, $T = t_n$, $\xi_i = \frac{n}{T}(t_i - t_{i-1})$, $g(\xi) = \{\xi f'(\xi)/f(\xi)\} + 1$, とおくと

$$\sum_{i=1}^n g(\xi_i) = 0, \quad \forall \xi_i > 0, \quad \sum \xi_i = n,$$

となる。いま $\xi_1 = \xi_2 = \dots = \xi_{n-1} = 1+h$, $\xi_n = 1-(n-1)h$,
 とおき $G(\xi) = g(1+\xi)$ とおくと

$$G(-(n-1)h) = -(n-1)G(h).$$

このことから $G(h) = -ch$ かつ $\int_0^\infty f(x) dx = 1$ に
 $\int_0^\infty x f(x) dx = 1$ により $c > 0$ に対して

$$f(x) = \frac{c^c}{\Gamma(c)} x^{c-1} e^{-cx}$$

が得られる。(かゝるに $t_n < T$ の場合には, この $f(x)$ に対
 し, $n = T$, $t = T - t_n$ とおくと

$$\frac{\partial}{\partial \theta} \log L_T\left(\frac{\eta}{T}\right) = ct + \frac{-f(t)}{1-F(t)} = 0$$

となつて定数 C は $C=1$ とする。よつて証明おひけり。

References

- [1] Kabanov, Yu. M. & Liptser, R. Sh. & Shiryaev, A. N., (1975), Martingale Method in the theory of Point Processes, Proceeding of Vilnius Symposium U.S.S.R. (in Russian).
- [2] Hawkes, A. G. & Oakes, D., (1974), A Cluster Process Representation of a Self-exciting Point Process, J. Appl. Prob., 11, 493-503.
- [3] Meyer, P. A., (1972), Martingale and Stochastic Integrals I, Lecture Notes in Mathematics, Springer-Verlag.
- [4] Ogata, Y. (1977), The Asymptotic Behavior of Maximum Likelihood Estimator of Stationary Point Processes, submitted to Annals of Institute of Statistical Mathematics.
- [5] Ozaki, T. (1977), The Maximum Likelihood Estimator of the Hawkes' Self-exciting Point Processes, This Volume.
- [6] Vere-Jones, D., (1975), Lectures on Point Processes, Department of Statistics, University of California, Berkeley.