Ann. Inst. Statist. Math.
37 (1985), Part A, 251-269

ON PROGRESSIVELY TRUNCATED MAXIMUM LIKELIHOOD ESTIMATORS

NOBUO INAGAKI* AND PRANAB KUMAR SEN**

(Received Jan. 5, 1984; revised Oct. 8, 1984)

Summary

The basic regularity conditions pertaining to the asymptotic theory of progressively truncated likelihood functions and maximum likelihood estimators are considered, and the uniform strong consistency and weak convergence of progressively truncated maximum likelihood estimators are studied systematically.

1. Introduction

In a clinical trial or life-testing experiment, observations (viz., failure times) are gathered, in order of magnitudes, sequentially over time. Further, due to cost, time and other limitations, often, experimentation is curtailed either after a pre-set duration of time (*truncation* or *type* I *censoring*) or after a prefixed number of responses occur (*type* II *censoring*). Interim analysis of accumulating data at various failure points (*progressive censoring*) or continuous statistical monitoring over the study period (*progressive truncation*) is sometimes advocated with a view to a possibly *early termination* of the study, contingent on the accumulating statistical evidence.

In a progressive censoring (PC) scheme, the *likelihood functions* of the accumulating data lead to a discrete time-parameter stochastic process (which, in general, may not have independent or stationary increments). Weak convergence of such a PC likelihood process, for the uni-parameter case, has been studied by Sen [9], and the same has been incorporated in study of the asymptotic properties of some *time-*

Key words and phrases: Progressive censoring, progressive truncation, maximum likelihood estimator, submartingale inequality, Gaussian process.

* Work done during the first author's visit (as a visiting scholar) to the University of North Carolina at Chapel Hill, supported by the Ministry of Education of the Japanese Government.

** Work supported by the (U.S.) National Heart, Lung and Blood Institute, Contact NIH-NHLBI-F1-2243-L.

sequential tests. For some related studies, we may also refer to Gardiner [4]. Under a progressive truncation (PT) scheme, one encounters a continuous time-parameter stochastic process relating to the accumulating likelihood functions and this also, generally, does not have homogeneous or independent increments. We may refer to Sen and Tsong [11], who studied the PT case and commented briefly on the multiparameter case too. Under more stringent regularity conditions and in a somewhat heuristic setup, such PC (or PT) likelihood functions have been employed by Bangdiwala [1] for some time-sequential tests for a real valued function of the parameter(s), granted some smoothness conditions. In this respect, the asymptotic theory of *progressively truncated maximum likelihood estimators* (PTMLE) provides the basic key for the study of the properties of the time-sequential tests, and there is substantial scope for more refined analysis.

For independent random variables (r.v.) or processes with independent increments, the intricate relationships between the likelihood function equations and maximum likelihood estimators have been studied by a host of workers; we may refer to Inagaki [5] where other references are also cited. In the context of PT or PC schemes, such relationships have not been studied yet in full generality. Indeed, in view of the lack of independence or homogeneity of the increments, such a study in a systematic and logically integrated manner may naturally require elaborate regularity conditions.

Section 2 is devoted to the details of this study. With a clear formulation of the martingale structure of PT likelihood processes, uniform continuity of the truncated Fisher information matrix and weak convergence of PT likelihood process in the general multiparameter case are studied in Section 3. Uniform strong consistency of the PTMLE is considered in Section 4. Weak convergence of PTMLE is studied in Section 5. This is primarily done through the exploration of the intricate relationships between the PT likelihood functions and the PTMLE's. The concluding section is devoted to some general remarks and discussion.

2. Preliminary notions and basic assumptions

Let $\{X_i: i\geqq 1\}$ be a sequence of independent and identically distributed (i.i.d.) r.v.'s with a distribution function (d.f.) $F_\theta(\cdot)$ and a probability density function (p.d.f.) $f_\theta(\cdot)$, both defined on $R^+=[0, \infty)$, where $\theta\in\Theta$, a subspace of $R^k=(-\infty, \infty)^k$, for some $k\geqq 1$. For every $n(\geqq 1)$, let $X_{n:1}\leqq\cdots\leqq X_{n:n}$ be the ordered r.v.'s corresponding to $X_1, \cdots, X_n$ (ties among these may be neglected, with probability 1). If the experiment involving n subjects is curtailed at a time point t (>0), the observable r.v.'s are then

$$(2.1)\qquad X_{n:j}\,,\quad j\leqq r_n(t)\,;\quad r_n(t)=\sum_{k=1}^{n} I(X_{n:k}\leqq t)\,,$$

where $X_{n:0}=0$ (conventionally) and $I(A)$ is the indicator function of the set A. For $0\leqq t<\infty$, let $\mathcal{B}_{n:t}=\mathcal{B}(X_{n:j},\ j\leqq r_n(t);\ t-X_{n:r_n(t)})$ be the σ-field generated by the observable r.v.'s at time t, with $\mathcal{B}_{n:0}=\mathcal{B}(\{\phi\})$ and $\mathcal{B}_{n:\infty}=\mathcal{B}(X_{n:j},\ j\leqq n)$. Also, let $\mathcal{F}_{n:t}$ be the space of $\mathcal{B}_{n:t}$-measurable functions. Then, for every n $(\geqq 1)$

$$(2.2)\qquad \mathcal{B}_{n:t} \text{ is nondecreasing in } t,\ t\in R^+\,.$$

Denote the projection mapping from $\mathcal{F}_{n:\infty}$ to $\mathcal{F}_{n:t}$ under the model distribution F_θ by $\mathcal{P}_{n:t}$: that is, for any integrable function $\psi(X_{n:1},\cdots, X_{n:n})$,

$$(2.3)\qquad (\mathcal{P}_{n:t}\psi)(X_{n:1},\cdots,X_{n:n})=\mathrm{E}_\theta\{\psi(X_{n:1},\cdots,X_{n:n})\,|\,\mathcal{B}_{n:t}\}\,,\qquad t\in R^+\,.$$

Let $\mathcal{B}_t=\mathcal{B}(X_i\leqq t,\ i\geqq 1)$, $t\in R^+$ with $\mathcal{B}_0=\mathcal{B}(\{\phi\})$ and $\mathcal{B}_\infty=\mathcal{B}(X_i,\ i\geqq 1)$, $\mathcal{F}_t$ be the space of $\mathcal{B}_t$-measurable functions and $\mathcal{P}_t$ be the projection mapping from $\mathcal{F}_\infty$ to $\mathcal{F}_t$. Then (2.2) and (2.3) hold for these as well. The usual likelihood function and equation for the full data $(X_1,\cdots, X_n)$ are

$$(2.4)\qquad L_n(\theta)=\prod_{i=1}^{n} f(X_i,\theta)\quad\text{or}\quad n!\prod_{i=1}^{n} f(X_{n:i},\theta)\,,$$

(where we write $f_\theta(x)=f(x,\theta)$) and

$$(2.5)\qquad \begin{aligned}\Phi_n(\theta)&=(\partial/\partial\theta)\log L_n(\theta)\\ &=\sum_{i=1}^{n}\phi(X_i,\theta)=0\,,\end{aligned}$$

respectively, where $\phi(x,\theta)=(\partial/\partial\theta)\log f(x,\theta)$ is the *likelihood score function* and $\Phi_n(\theta)$ is the *likelihood estimating function.* Let us define the *survival function* $\bar{F}_\theta(t)$ and the *survival likelihood score function* $\bar{\phi}(t,\theta)$, $t\in R^+$, by

$$(2.6)\qquad \bar{F}_\theta(t)=1-F_\theta(t)=\int_t^\infty f(x,\theta)dx\,,$$

$$(2.7)\qquad \begin{aligned}\bar{\phi}(t,\theta)&=(\partial/\partial\theta)\log\bar{F}_\theta(t)\\ &=\int_t^\infty \phi(x,\theta)F_\theta(dx)/\bar{F}_\theta(t)\,,\end{aligned}$$

where (2.7) holds under some regularity conditions, which will be introduced in the sequel. We introduce the projection likelihood function and the projection likelihood score function under F_θ (as in [11]) by letting

$$(2.8)\qquad f_t(x, \theta)=\mathrm{E}_\mu\,[f(X_1, \theta)|\mathcal{B}_{1:t}]\overset{*}{=}\begin{cases} f(x, \theta)\,, & \text{if } 0\leqq x\leqq t\,,\\ \bar{F}_\theta(t)\,, & \text{if } t<x<\infty\,,\end{cases}$$

and

$$(2.9)\qquad \phi_t(x, \theta)=\mathrm{E}_\theta\,(\phi(X_1, \theta)|\mathcal{B}_{1:t}]=\begin{cases} \phi(x, \theta)\,, & \text{if } 0\leqq x\leqq t\,,\\ \bar{\phi}(t, \theta)\,, & \text{if } t<x<\infty\,.\end{cases}$$

$\mathrm{E}_\mu\,[f|\mathcal{B}_{1:t}]$ stands for the generalized conditional expectation (c.f. Rao [8]) and the equality ($\overset{*}{=}$) means the following: Let $B_t=B_\frown[0, t]$.

$$(2.10)\qquad \int_B f_t(x, \theta)dx=\int_B f(x, \theta)dx\,, \qquad \text{for any } B\in\mathcal{B}_{1:t}\,.$$

$$\overset{*}{=}\begin{cases} \displaystyle\int_{B_t} f(x, \theta)dx\,, & \text{if } B=B_t\,,\\ \displaystyle\int_{B_t} f(x, \theta)dx+\bar{F}_\theta(t)\,, & \text{if } B=B_t\smile(t, \infty)\,.\end{cases}$$

Let us now introduce the basic regularity assumptions on which the proposed theory rests. Suppose that the true parameter θ_0 is any inner point of Θ (but fixed). Let $\|\cdot\|$ be the Euclidean norm.

[A1] The parameter space Θ is a subspace of R^k and for any $M>0$, Θ_M $(=\Theta_\frown\{\theta\colon \|\theta\|\leqq M\})$ is closed.

[A2] For each $\theta\in\Theta$, $F_\theta(x)$ has a positive p.d.f. $f(x, \theta)$ (with respect to the Lebesgue measure), where $f(x, \theta)$ is continuous in Θ for all $x\in R^+$.

[A3] If $\theta_1\neq\theta_2$, $\int_{R^+}|f(x, \theta_1)-f(x, \theta_2)|dx>0$. Further, if $\theta_1\neq\theta_2$ and $F_{\theta_1}(t)=F_{\theta_2}(t)$ for some $t>0$, then

$$(2.11)\qquad \int_0^t|f(x, \theta_1)-f(x, \theta_2)|dx>0\,.$$

In addition to [A1]–[A3], we have the following ones, where $U_d(\theta_0)=\{\theta\colon \|\theta-\theta_0\|<d\}$ stands for a neighborhood of θ_0 of radius d (>0).

[B1] There exists a $d_0>0$, such that for every $x\in R^+$, $f(x, \theta)$ is continuously twice differentiable in $\theta\in U_{d_0}(\theta_0)$ and there are positive functions $U_i^*(x)=U_i^*(x;\theta_0, d_0)$ with

$$(2.12)\qquad U_i^*=\int_{R^+}U_i^*(x)dx<\infty\,, \qquad i=1, 2\,,$$

such that for every $\theta\in U_{d_0}(\theta_0)$,

$$(2.13)\qquad \|(\partial/\partial\theta)f(x, \theta)\|\leqq U_1^*(x)\,, \qquad x\in R^+\,,$$

$$(2.14)\qquad \|(\partial^2/\partial\theta\partial\theta')f(x, \theta)\|\leqq U_2^*(x)\,, \qquad x\in R^+\,.$$

Define $\phi(x, \theta)$ as in after (2.5) and let

(2.15) $$\dot{\phi}(x, \theta)=(\partial/\partial\theta')\phi(x, \theta)=(\partial^2/\partial\theta\partial\theta')\log f(x, \theta)\ .$$

[B2] The Fisher information matrix

(2.16) $$I(\theta)=\mathrm{E}_\theta\{[\phi(X_i, \theta)][\phi(X_i, \theta)]'\}$$

exists, is positive definite (p.d.) and continuous in $\theta \in U_{d_0}(\theta_0)$. The expectation matrix $\mathrm{E}_\theta\dot{\phi}(X_i, \theta)$ exists and is continuous in $\theta \in U_{d_0}(\theta_0)$. $\mathrm{E}_{\theta_0}\|\dot{\phi}(X_i, \theta_0)\|^2=\mathrm{E}_{\theta_0}\{\mathrm{trace}\,(\dot{\phi}(X_i, \theta_0))^2\}$ exists.

[B3] Let, for every $d>0$,

(2.17) $$u(x; \theta_0, d)=\sup\{\|\dot{\phi}(x, \theta)-\dot{\phi}(x, \theta_0)\|;\ \|\theta-\theta_0\|<d\}\ .$$

Then, the expected value of $u(X_i; \theta_0, d)$ exists and

(2.18) $$\lim_{d\downarrow 0}\mathrm{E}_{\theta_0}u(X_i; \theta_0, d)=0\ .$$

[B4] For every d and $T>0$, let

(2.19) $$\bar{R}_T(d)=\sup\{\bar{F}_{\theta_0}(t)\|(\partial^2/\partial\theta\partial\theta')\log\bar{F}_\theta(t)-(\partial^2/\partial\theta\partial\theta')\log\bar{F}_{\theta_0}(t)\|:\ \|\theta-\theta_0\|<d,\ t>T\}\ .$$

Then, for all sufficiently small $d>0$,

(2.20) $$\lim_{T\to\infty}\bar{R}_T(d)=0\ .$$

For some other regularity conditions, let's introduce the following:

(2.21) $$g(x; \theta, d)=\sup\{f(x, \tau):\ \tau\in\Theta,\ \|\tau-\theta\|<d\}\ ,\qquad d>0\ ,$$

(2.22) $$\tilde{G}_\theta(t; d)=\sup\{\bar{F}_\tau(t):\ \tau\in\Theta,\ \|\tau-\theta\|<d\}\ ,\qquad d>0\ ,$$

(2.23) $$\tilde{g}_t(x; \theta, d)=\begin{cases} g(x; \theta, d)\ , & \text{if } x\in[0, t]\ ,\\ \tilde{G}_\theta(t; d)\ , & \text{if } x\in(t, \infty),\ t\in R^+\ .\end{cases}$$

Furthermore, let

(2.24) $$g(x; \theta_\infty, d)=\sup\{f(x, \theta):\ \theta\in\Theta,\ \|\theta-\theta_0\|>d^{-1}\}\ ,$$

(2.25) $$\tilde{G}_{\theta_\infty}(t; d)=\sup\{\bar{F}_\theta(t):\ \theta\in\Theta,\ \|\theta-\theta_0\|>d^{-1}\}\ ,$$

(2.26) $$\tilde{g}_t(x; \theta_\infty, d)=\begin{cases} g(x; \theta_\infty, d)\ , & \text{if } x\in[0, t]\ ,\\ \tilde{G}_{\theta_\infty}(t; d)\ , & \text{if } x\in(t, \infty),\ t\in R^+\ .\end{cases}$$

[C1] For every $\theta\in\Theta$, there exists a positive number $d_1(\theta)>0$ such that, for any $d\in(0, d_1(\theta))$, (2.27) $$\int_0^\infty g(x;\theta,d)dx<\infty\ .$$

[C2] There exists a positive number $\alpha_1>0$ such that

(2.28) $$\lim_{d\downarrow 0}\int_0^\infty \{g(x;\theta_\infty,d)/f(x,\theta_0)\}^{\alpha_1}f(x,\theta_0)dx=0\ .$$

Furthermore, for all sufficiently small $d>0$ and any $t>0$,

(2.29) $$\int_0^\infty \log^+\{g(x;\theta_\infty,d)/f(x,\theta_0)\}f(x,\theta_0)dx<\infty\ ,$$

(2.30) $$\int_0^t \log^-\{g(x;\theta_\infty,d)/f(x,\theta_0)\}f(x,\theta_0)dx>0\ , \qquad \text{and}$$

(2.31) $$\lim_{T\uparrow\infty}\int_T^\infty \log^-\{g(x;\theta_\infty,d)/f(x,\theta_0)\}f(x,\theta_0)dx/\bar{F}_{\theta_0}(T)<\infty\ ,$$

where $\log^+ x=\max\{\log x, 0\}$ and $\log^- x=\max\{-\log x, 0\}$.

The following remarks regarding the three sets of assumptions are worth mentioning.

(I) Assumption [B1] implies that $f(x,\theta)$ and $(\partial/\partial\theta)f(x,\theta)$ are *differentiable in mean*, i.e.

(2.32) $$\lim_{\|h\|\to 0}\int_0^\infty |f(x,\theta+h)-f(x,\theta)-(\partial/\partial\theta)f(x,\theta)'h|dx/\|h\|=0\ ,$$

(2.33) $$\lim_{\|h\|\to 0}\frac{1}{\|h\|}\int_0^\infty \|(\partial/\partial\theta)f(x,\theta+h)-(\partial/\partial\theta)f(x,\theta)-(\partial^2/\partial\theta\partial\theta')f(x,\theta)h\|dx =0\ .$$

These, in turn, lead to

(2.34) $$\int_0^\infty (\partial/\partial\theta)f(x,\theta)dx=0\ ,$$

(2.35) $$\int_0^\infty (\partial^2/\partial\theta\partial\theta')f(x,\theta)dx=0\ .$$

It also follows from assumptions [B1] and [B3] that

(2.36) $$\lim_{\|h\|\to 0} \mathrm{E}_\theta\|\phi(X_i,\theta+h)-\phi(X_i,\phi)-\dot{\phi}(X_i,\theta)h\|/\|h\|=0\ ,$$

so that, by (2.35) and (2.36), we have

(2.37) $$\mathrm{E}_\theta\,\dot{\phi}(X_i,\theta)=-I(\theta)\ .$$

If we define for every $t\in R^+$, $\theta\in\boldsymbol{\Theta}$,

(2.38) $$I_t(\theta)=\mathrm{E}_\theta\{[\phi_t(X_i,\theta)][\phi_t(X_i,\theta)]'\}\ ,$$

then, by (2.9), (2.16) and (2.38),

$$I_t(\theta)=I(\theta)-\bar{I}_t(\theta)\,,\qquad t\in R^+\,, \tag{2.39}$$

where

$$\begin{aligned}\bar{I}_t(\theta)&=\int_t^\infty \{\phi(x,\theta)-\bar{\phi}(t,\theta)\}\{\phi(x,\theta)-\bar{\phi}(t,\theta)\}'F_\theta(dx)\\&=\int_t^\infty (\phi(x,\theta))(\phi(x,\theta))'F_\theta(dx)-\bar{F}_\theta(t)(\bar{\phi}(t,\theta))(\bar{\phi}(t,\theta))'\,.\end{aligned} \tag{2.40}$$

That is, $I_t(\theta)$ is equal to $J_{F_\theta(t)}(\theta)$ in [11]. Note that $I_\infty(\theta)=\bar{I}_0(\theta)=I(\theta)$ and $I_0(\theta)=\bar{I}_\infty(\theta)=0$. From Assumption [B1], (2.36) and (2.39), we have

$$\begin{aligned}(\partial^2/\partial\theta\partial\theta')\log\bar{F}_\theta(t)&=(\partial/\partial\theta)\bar{\phi}(t,\theta)\\&=\int_t^\infty (\partial^2/\partial\theta\partial\theta')f(x,\theta)dx/\bar{F}_\theta(t)\\&\quad-\left\{\int_t(\partial/\partial\theta)f(x,\theta)dx\right\}\left\{\int_t(\partial/\partial\theta)f(x,\theta)dx\right\}'\Big/\bar{F}_\theta^2(t)\\&=\left\{\int_t^\infty \dot{\phi}(x,\theta)f(x,\theta)dx\right\}\Big/\bar{F}_\theta(t)+\bar{I}_t(\theta)/\bar{F}_\theta(t)\,.\end{aligned} \tag{2.41}$$

(See Theorem 2.1 and 2.2 of [6] in this context.)

(II) In the same way as in (I), [B1] implies that

$$\begin{aligned}\Delta_i(d)=\sup&\left\{\int_0^\infty \|(\partial^i/\partial\theta^i)f(x,\theta)-(\partial^i/\partial\theta^i)f(x,\theta_0)\|dx:\ \|\theta-\theta_0\|<d\right\}\\&\to 0 \text{ as } d\downarrow 0\,,\qquad \text{for } i=1,2\,,\end{aligned} \tag{2.42}$$

where $(\partial^2/\partial\theta^2)f$ stands for the general term of $(\partial^2/\partial\theta\partial\theta')f$.

(III) Assumption [C1] holds if [B1] holds for every $\theta_0\in\Theta$.

3. Some preliminary results

From (2.8) and (2.9), we have the PT likelihood function and equation for the set of observations $\{X_i\leq t,\ 1\leq i\leq n\}$ as follows:

$$\begin{aligned}L_{n:t}(\theta)&=\prod_{i=1}^n f_t(X_i,\theta)\\&\left(\text{or}\ =\binom{n}{r_n(t)}(r_n(t))!\left\{\prod_{i=1}^{r_n(t)}f(X_{n:i},\theta)\right\}\{\bar{F}_\theta(t)\}^{n-r_n(t)}\right),\end{aligned} \tag{3.1}$$

$$\begin{aligned}\Phi_{n:t}(\theta)&=\sum_{i=1}^n \phi_t(X_i,\theta)\\&=\sum_{i=1}^{r_n(t)}\phi(X_{n:i},\theta)+(n-r_n(t))\bar{\phi}(t,\theta)=0\,,\end{aligned} \tag{3.2}$$

where $\phi(X_{n:0},\theta)=0$. Note that $L_{n:0}(\theta)=1$, $\Phi_{n:0}(\theta)=0$; $L_{n:\infty}(\theta)=L_n(\theta)$ and

$\Phi_{n:\infty}(\theta)=\Phi_n(\theta)$. The facts that

$$L_{n:t}(\theta)=\mathcal{P}_{n:t}L_n(\theta)=\mathrm{E}_\mu\,[L_n(\theta)\,|\,\mathcal{B}_{n:t}]\,, \tag{3.3}$$

$$\Phi_{n:t}(\theta)=\mathcal{P}_{n:t}\Phi_n(\theta)=\mathrm{E}_\theta\,[\Phi_n(\theta)\,|\,\mathcal{B}_{n:t}]\,, \tag{3.4}$$

for $t\in[0,\infty)$ (viz. [11]) play a vital role in the developments of this paper. Since, from (2.9) and (2.38),

$$\mathrm{E}_\theta\,\phi_t(X_i,\theta)=\mathrm{E}_\theta\,\phi(X_i,\theta)=0\,, \tag{3.5}$$

$$V_\theta\,\phi_t(X_i,\theta)=I_t(\theta)=I(\theta)-\bar{I}_t(\theta)\,, \tag{3.6}$$

for each $t\in R^+$, the sample functions $\Phi_n^\circ=\{\Phi_{n:t}^\circ(\theta_0);\ t\in R^+\}$ with

$$\begin{aligned}\Phi_{n:t}^\circ(\theta_0)&=n^{-1/2}\Phi_{n:t}(\theta_0)=n^{-1/2}\sum_{i=1}^n\phi_t(X_i,\theta_0)\\&=n^{-1/2}\{\mathcal{P}_{n:t}\Phi_n(\theta_0)\}\end{aligned} \tag{3.7}$$

has the following moment structures under P_{θ_0}:

$$\mathrm{E}_{\theta_0}\{\Phi_{n:t}^\circ(\theta_0)\}=0\,, \tag{3.8}$$

$$\mathrm{E}_{\theta_0}\{\Phi_{n:s}^\circ(\theta_0)\Phi_{n:t}^\circ(\theta_0)'\}=I_{s\wedge t}(\theta_0)\,, \tag{3.9}$$

for every $s,t\in R^+$, where $s\wedge t=\text{minimum}\,(s,t)$. Furthermore, under P_{θ_0}, $\{\Phi_{n:t}^\circ(\theta_0),\mathcal{B}_{n:t};\ t\in R^+\}$ (for every $n\geqq 1$) is a zero mean martingale closed on the right by $n^{-1/2}\Phi_n(\theta_0)$ by virtue of (3.4) and (3.7).

Now, let $W^\circ=\{W^\circ(t),\ 0\leqq t\leqq 1\}$ be a Brownian bridge on [0, 1] (so that W° is Gaussian with $\mathrm{E}\{W^\circ\}=0$ and $\mathrm{E}\{W^\circ(s)W^\circ(t)\}=s\wedge t-st$ for each $s,t\in[0,1]$). We consider a k-variate Gaussian process $\Phi^\circ=\{\Phi_t^\circ(\theta_0),\ t\in R^+\}$, where

$$\Phi_t^\circ(\theta_0)=\int_0^\infty\phi_t(x,\theta_0)\,W^\circ(F_{\theta_0}(dx))\,; \tag{3.10}$$

$$\Phi_\infty^\circ(\theta_0)=\Phi^\circ(\theta_0)=\int_0^\infty\phi(x,\theta_0)\,W^\circ(F_{\theta_0}(dx))\,, \tag{3.11}$$

and $\Phi_0^\circ(\theta_0)=0$, with probability 1. Then, as in after (2.3),

$$\Phi_t^\circ(\theta_0)=\mathcal{P}_t\Phi(\theta_0)=\mathrm{E}_{\theta_0}\,[\Phi^\circ(\theta_0)\,|\,\mathcal{B}_t]\,,\qquad t\in R^+\,, \tag{3.12}$$

so that $\Phi^\circ=\{\Phi_t^\circ(\theta_0),\mathcal{B}_t;\ t\in R^+\}$ is a martingale closed on the right by $\Phi_\infty^\circ(\theta_0)$ and has the same moment structures as Φ_n° defined in (3.7):

$$\mathrm{E}_{\theta_0}\{\Phi_t^\circ(\theta_0)\}=0\,, \tag{3.13}$$

$$\mathrm{E}_{\theta_0}\{\Phi_s^\circ(\theta_0)\Phi_t^\circ(\theta_0)'\}=I_{s\wedge t}(\theta_0)\,, \tag{3.14}$$

for each $s,t\in R^+$. Furthermore, we have the following.

THEOREM 3.1. *Suppose the set of assumptions, A and B, in Section 2. Then, under* P_{θ_0}, $\Phi_n^\circ = \{\Phi_{n:t}^\circ(\theta_0);\ t \in R^+\}$ *converges weakly to the Gaussian process* $\Phi^\circ = \{\Phi_t^\circ(\theta_0);\ t \in R^+\}$, *in the extended Skorokhod* J_1*-topology on* $D^k[R^+]$.

This theorem is essentially due to Sen and Tsong [11]. The convergence of the finite-dimensional distributions (f.d.d.) of $\{\Phi_n^\circ\}$ to those of Φ° follows from the classical central limit theorem. Note that by construction both processes belong to $D^k[R^+]$. The proof of the tightness of the former processes is parallel to that due to Sen and Tsong [11] by using Lemma 2.1 and the submartingale inequality of Lemma 3.3 of [11].

The following lemma is easy to see.

LEMMA 3.2. *Under the regularity conditions of Section* 2,

(i) $I_t(\theta)$, $t \in R^+$, $\theta \in \Theta$, *is continuous in* $R^k \times U_d(\theta_0)$,

(ii) $H(d) = \sup\{\|I_t(\theta) - I_t(\theta_0)\| :\ \|\theta - \theta_0\| < d,\ t \in R^+\} \to 0$, *as* $d \downarrow 0$,

(iii) $\bar{H}(d) = \sup\{\|\bar{I}_t(\theta) - \bar{I}_t(\theta_0)\| :\ \|\theta - \theta_0\| < d,\ t \in R^+\} \to 0$, *as* $d \downarrow 0$.

(iv) $$Y(d) = \sup\left\{\left\|\int_t^\infty [\phi(x, \theta)][\phi(x, \theta)]' f(x, \theta)dx - \int_t^\infty [\phi(x, \theta_0)][\phi(x, \theta_0)]' f(x, \theta_0)dx\right\| :\ \|\theta - \theta_0\| < d,\ t \in R^+\right\} \to 0,$$ *as* $d \downarrow 0$.

4. Uniform strong consistency of PTMLE

Let us define the PTMLE $\{\hat{\theta}_{n:t}, t \in R^+\}$ by the following equation,

$$L_{n:t}(\hat{\theta}_{n:t}) = \sup\{L_{n:t}(\theta) :\ \theta \in \Theta\}. \tag{4.1}$$

Since $L_{n:t}(\theta)$ is continuous in $\theta \in \Theta$ (for every $n \geqq 1$ and $t \in R^+$) and further $L_{n:t}(\theta)$ is right-continuous in $t \in R^+$ (for every n and θ), we claim that for every n, $\hat{\theta}_{n:t}$ is right-continuous in $t \in R^+$: that is,

$$\{\hat{\theta}_{n:t}, t \in R^+\} \in D^k[R^+], \qquad \text{for any } n \geqq 1. \tag{4.2}$$

The main result of this section relates to the uniform strong consistency of $\hat{\theta}_{n:t}$ and towards this, we consider first following.

LEMMA 4.1. *Under Assumptions A and* [C1] *in Section* 2, *for every positive number* d, M *and* t_1, *there exists a* ρ_0: $0 < \rho_0 < 1$, *such that for any* $\eta > 0$ *and all* n *sufficiently large*

$$P_{\theta_0}[\sup\{L_{n:t}(\theta)/L_{n:t}(\theta_0) :\ t \in [t_1, \infty),\ d \leqq \|\theta - \theta_0\| \leqq M\} > \eta] \leqq \rho_0^n. \tag{4.3}$$

PROOF. Define $g(x;\theta,d)$ and $\tilde{g}_t(x;\theta,d)$ as in (2.21) and (2.23) and let the projection of g be

$$g_t(x;\theta,d)=\mathrm{E}_\mu[g(X_1;\theta,d)|\mathcal{B}_{1;t}]\overset{*}{=}\begin{cases} g(x;\theta,d), & \text{if } 0\leqq x\leqq t, \\ \bar{G}_\theta(t;d), & \text{if } t<x<\infty, \end{cases} \tag{4.4}$$

where

$$\bar{G}_\theta(t;d)=\int_t^\infty g(x;\theta,d)dx, \quad t\in R^+, \quad d>0, \quad \theta\in\Theta. \tag{4.5}$$

Then, by (2.21)–(2.23) and (4.4)–(4.5), we have

$$f_t(x,\theta)\leqq\tilde{g}_t(x;\theta,d)\leqq g_t(x;\theta,d), \qquad \text{for any } t\in R^+; \tag{4.6}$$

$$\lim_{d\downarrow 0} g_t(x;\theta,d)=f_t(x,\theta). \tag{4.7}$$

Also, by the Lebesgue dominated convergence theorem, we have from Assumption [A3]

$$\begin{aligned}\lim_{d\downarrow 0}\mathrm{E}_{\theta_0}[\log\{g_t(X_i;\theta,d)/f_t(X_i,\theta_0)\}]&=\mathrm{E}_{\theta_0}[\log\{f_t(X_i,\theta)/f_t(X_i,\theta_0)\}]\\ &=-K_t(\theta,\theta_0) \quad \text{(say)},\end{aligned} \tag{4.8}$$

where, for every $t>0$,

$$K_t(\theta,\theta_0)>0, \qquad \text{if } \theta\neq\theta_0. \tag{4.9}$$

Therefore, from (4.8) and (4.9), we claim that for every $t_1>0$ there exists a positive number $d_2(\theta)>0$ such that for any $d\in(0,d_2(\theta))$

$$\mathrm{E}_{\theta_0}[\log\{g_{t_1}(X_i;\theta,d)/f_{t_1}(X_i,\theta_0)\}]\leqq-\frac{1}{2}K_{t_1}(\theta,\theta_0)<0. \tag{4.10}$$

Let us next define Θ_M as in [A1] in Section 2 and $U_d(\theta_0)$ as in after (2.10). Let then $\Theta_{M,d}=\Theta_M-U_d(\theta_0)$, $d>0$. Then, by Assumption [C1] and (4.10), we have that for every $d>0$, $M>0$ and $t_1>0$, there exist finite points $\theta_1,\cdots,\theta_m$ and finite numbers $d_1,\cdots,d_m$ with $0<d_j<\min(d_1(\theta_j),d_2(\theta_j))$, $1\leqq j\leqq m$, (for $d_1(\theta)$ in Assumption [C1] and $d_2(\theta)$ above) such that

$$\Theta_{M,d}<\bigcup_{j=1}^m U_{d_j}(\theta_j), \tag{4.11}$$

$$\int_0^\infty g(x;\theta_j,d_j)dx<\infty, \tag{4.12}$$

and, for each j $(=1,\cdots,m)$,

$$\mathrm{E}_{\theta_0}[\log\{g_{t_1}(X_i;\theta_j,d_j)/f_{t_1}(X_i,\theta_0)\}]\leqq-\frac{1}{2}K_{t_1}(\theta_j,\theta_0)<0. \tag{4.13}$$

Set

$$K = \min\{K_{t_1}(\theta_j, \theta_0);\ j=1, \cdots, m\}. \tag{4.14}$$

Then, from (2.21)–(2.25), (4.7) and (4.11), we obtain that for every $t_1 > 0$,

$$\begin{aligned} &P_{\theta_0}\{\sup[L_{n:t}(\theta)/L_{n:t}(\theta_0):\ t_1 \leqq t < \infty,\ \theta \in \Theta_{M,d}] > e^{-nK}\} \\ &\leqq \sum_{j=1}^{m} P_{\theta_0}\left\{\sup\left[\prod_{i=1}^{n}(\tilde{g}_t(X_i;\theta_j, d_j)/f_t(X_i, \theta_0));\ t_1 \leqq t < \infty\right] > e^{-nK}\right\} \\ &\leqq \sum_{j=1}^{m} P_{\theta_0}\left\{\sup\left[\prod_{i=1}^{n}(g_t(X_i;\theta_j, d_j)/f_t(X_i, \theta_0));\ t_1 \leqq t < \infty\right] > e^{-nK}\right\}. \end{aligned} \tag{4.15}$$

Next, we note that $\left\{\prod_{i=1}^{n}(g_t(X_i;\theta_j, d_j)/f_t(X_i, \theta_0)),\ \mathcal{B}_{n:t};\ t \in R^+\right\}$ is a martingale for every n, θ_j and d_j, and hence

$$\left\{\prod_{i=1}^{n}(g_t(X_i;\theta_j, d_j)/f_t(X_i, \theta_0))^{\alpha},\ \mathcal{B}_{n:t};\ t \in R^+\right\} \text{ is a super-martingale for every } \alpha:\ 0 < \alpha \leqq 1. \tag{4.16}$$

Therefore, by the super-martingale inequality (viz. [7], p. 281), we obtain that for any α: $0 < \alpha \leqq 1$, $t_1 > 0$

$$\begin{aligned} &P_{\theta_0}\left\{\sup\left[\prod_{i=1}^{n}(g_t(X_i;\theta_j, d_j)/f_t(X_i, \theta_0)):\ t_1 \leqq t < \infty\right] > e^{-nK}\right\} \\ &\leqq e^{n\alpha K}\,\mathrm{E}_{\theta_0}\left[\prod_{i=1}^{n}(g_{t_1}(X_i;\theta_j, d_j)/f_{t_1}(X_i, \theta_0))\right]^{\alpha} \\ &= \{e^{\alpha K}\,\mathrm{E}_{\theta_0}[g_{t_1}(X_1;\theta_j, d_j)/f_{t_1}(X_1, \theta_0)]^{\alpha}\}^n. \end{aligned} \tag{4.17}$$

By the Chernoff information inequality ([2]), we have from (4.13) and (4.14) that

$$\rho(\theta_j) = \inf\{e^{\alpha K}\,\mathrm{E}_{\theta_0}[g_{t_1}(X_1;\theta_j, d_j)/f_{t_1}(X_1, \theta_0)]^{\alpha}:\ 0 < \alpha \leqq 1\} < 1, \tag{4.18}$$

for every j $(=1, \cdots, m)$. Corresponding to the η in (4.3), we let $n_0 \geqq -K\log\eta$ and let

$$\rho_0 = m^{1/n_0}\max\{\rho(\theta_j);\ 1 \leqq j \leqq m\} < 1. \tag{4.19}$$

Then, from (4.15) and (4.17)–(4.19), we conclude that for every $n \geqq n_0$

$$\begin{aligned} &P_{\theta_0}\{\sup[L_{n:t}(\theta)/L_{n:t}(\theta_0):\ t_1 \leqq t < \infty,\ \theta \in \Theta_{M,d}] > \eta\} \\ &\leqq P_{\theta_0}\{\sup[L_{n:t}(\theta)/L_{n:t}(\theta_0):\ t_1 \leqq t < \infty,\ \theta \in \Theta_{M,d}] > e^{-nK}\} \\ &\leqq \sum_{j=1}^{m}[\rho(\theta_j)]^n \\ &\leqq m[\max\{\rho(\theta_j);\ 1 \leqq j \leqq m\}]^n \leqq \rho_0^n, \end{aligned} \tag{4.20}$$

and this completes the proof of (4.3).

LEMMA 4.2. *Under Assumption* [C2], *for every* $t_1>0$, *there exist positive numbers* $M_1>0$ *and* $0<\rho_1<1$ *such that for any* η, $0<\eta<1$ *and all* n *sufficiently large,*

$$P_{\theta_0}\{\sup\,[L_{n:t}(\theta)/L_{n:t}(\theta_0):\ t_1\leqq t<\infty,\ \|\theta-\theta_0\|>M_1]>\eta\}\leqq\rho_1^n\,. \tag{4.21}$$

PROOF. Let $d=M_1^{-1}$. From (2.21)-(2.23), it follows that

$$\begin{aligned}\sup\,&\{[L_{n:t}(\theta)/L_{n:t}(\theta_0)]^{\alpha_1}:\ \|\theta-\theta_0\|>M_1\}\\ &\leqq\prod_{i=1}^{n}\{\tilde g_t(X_i;\,\theta_\infty,\,d)/f_t(X_i,\,\theta_0)\}^{\alpha_1}\\ &=\exp\Big[\sum_{i=1}^{r_n(t)}\log\{g(X_{n:i};\,\theta_\infty,\,d)/f(X_{n:i},\,\theta_0)\}^{\alpha_1}\\ &\qquad+n\bar F_n(t)\log\{\tilde G_{\theta_\infty}(t;\,d)/\bar F_{\theta_0}(t)\}^{\alpha_1}\Big]\,.\end{aligned} \tag{4.22}$$

Let $h(x;d)=\log\{g(x;\,\theta_\infty,\,d)/f(x,\,\theta_0)\}^{\alpha_1}$ and consider its projection:

$$h_t(x;d)=\mathrm{E}_{\theta_0}[h(X_1;d)\,|\,\mathcal{B}_{1:t}]=\begin{cases}h(x;d)\,, & \text{if } x\in[0,t]\,,\\ \displaystyle\int_t^\infty h(x;d)f(x,\theta_0)dx/\bar F_{\theta_0}(t)\,, & \\ & \text{if } x\in(t,\infty)\,.\end{cases} \tag{4.23}$$

Then, the last exponential value of (4.22) is bounded from above by

$$\begin{aligned}\exp\Big\{&\sum_{i=1}^{n}h_t(X_i;d)\\ &+n\bar F_n(t)\int_t^\infty h^-(x;d)f(x,\theta_0)dx/\bar F_{\theta_0}(t)+n\bar F_n(t)\log F_{\theta_0}^{-1}(t)\Big\}\,,\end{aligned} \tag{4.24}$$

because of $\tilde G_{\theta_\infty}(t;d)\leqq 1$.

Now, it is easy to see that for every n,

$$\{\exp[n\bar F_n(t)\log\bar F^{-1}(t)],\ \mathcal{B}_{n:t};\ t\in R^+\}\ \text{is a submartingale.} \tag{4.25}$$

Furthermore, for every n,

$$\mathrm{E}_{\theta_0}\{\exp[n\bar F_n(t)\log\bar F_{\theta_0}^{-1}(t)]\}=\{2-\bar F_{\theta_0}(t)\}^n\leqq 2^n\,. \tag{4.26}$$

Thus, by the submartingale inequality (viz. [10], p. 16), we have that for every n and $M_2>\log 2$,

$$P_{\theta_0}\{\sup\,[\exp(n\bar F_n(t)\log\bar F^{-1}(t));\ 0\leqq t<\infty]\geqq e^{nM_2}\}\leqq\rho_2^n\,, \tag{4.27}$$

where $0<\rho_2=2e^{-M_2}<1$.

Next, (2.31) in Assumption [C2] implies that there exist positive numbers M_3 and T such that

$$\int_t^\infty h^-(x;d)f(x,\theta_0)dx/\bar{F}_{\theta_0}(t)<M_3\,,\qquad \text{for every } t>T\,. \tag{4.28}$$

Thus, for every n and all sufficiently small $d>0$,

$$\sup_{T\le t<\infty}\left\{\exp\left[n\bar{F}_n(t)\int_t^\infty h^-(x;d)f(x,\theta_0)dx/\bar{F}_{\theta_0}(t)\right]\right\}<e^{nM_3}\,,\quad \text{a.s..} \tag{4.29}$$

Since $\left\{\exp\left[n(\bar{F}_n(t)/\bar{F}_{\theta_0}(t)-1)\int_{t_1}^\infty h^-(x;d)f(x,\theta_0)dx\right],\ \mathscr{B}_{n:t};\ t_1\le t\le T\right\}$ is easily seen to be a submartingale for every n and $t_1>0$, we have by the submartingale inequality that for any $\varepsilon>0$ and $\alpha>0$

$$\begin{aligned} P_{\theta_0}&\left\{\sup_{t_1\le t\le T}\exp\,[n(\bar{F}_n(t)/\bar{F}_{\theta_0}(t)-1)M_4]>e^{n}\right\} \\ &\le e^{-\alpha n\varepsilon}\,\mathrm{E}_{\theta_0}\{\exp\,[\alpha n(\bar{F}_n(T)/\bar{F}_{\theta_0}(T)-1)M_4]\} \\ &=\{e^{-\alpha(\varepsilon+M_4)}\bar{F}_{\theta_0}(T)\exp\,[\alpha M_4/\bar{F}_{\theta_0}(T)]\}^n\,, \end{aligned} \tag{4.30}$$

letting $M_4=M_4(d)=\int_{t_1}^\infty h^-(x;d)f(x,\theta_0)dx$. By the Chernoff's information inequality, it is shown that

$$\rho_3=\inf_{\alpha>0}\{e^{-\alpha(\varepsilon+M_4)}\bar{F}_{\theta_0}(T)\exp\,[\alpha M_4/\bar{F}_{\theta_0}(T)]\}<1\,. \tag{4.31}$$

Therefore, this and (4.29) conclude that for $M_3'=\max\,(M_3,\,M_4+\varepsilon)$,

$$P_{\theta_0}\left\{\sup_{t_1\le t<\infty}\exp\left[n\bar{F}_n(t)\int_t^\infty h^-(x;d)f(x,\theta_0)dx/\bar{F}_{\theta_0}(t)\right]>e^{nM_3'}\right\}<\rho_3^n\,. \tag{4.32}$$

On the other hand, it follows from (2.28) in Assumption [C2] that

$$\begin{aligned} \mathrm{E}_{\theta_0}h_t(X_i;d)&=\mathrm{E}_{\theta_0}h(X_i;d) \\ &\le\log\int_0^\infty\{g(x;\theta_\infty,d)/f(x,\theta_0)\}^{\alpha_1}f(x,\theta_0)dx\to-\infty\,, \\ &\qquad\qquad\qquad\text{as } d\to0\,, \end{aligned}$$

and hence, from (2.29) and (2.30) that for all sufficiently small $d>0$ and some $\varepsilon_1>0$

$$\mathrm{E}_{\theta_0}h(X_i;d)+M_2+M_3'<-\varepsilon_1<0\,. \tag{4.33}$$

Since $\left\{\sum_{i=1}^n h_t(X_i;d),\ \mathscr{B}_{n:t};\ t\in R^+\right\}$ is a martingale for every n, we obtain by the submartingale inequality that for any α, $0<\alpha<1$,

$$\begin{aligned} P_{\theta_0}&\left\{\sup_{t_1\le t<\infty}\exp\left[\sum_{i=1}^n h_t(X_i;d)+nM_2+nM_3'\right]>e^{-n\varepsilon_1}\right\} \\ &\le\{\exp\,(\alpha(\varepsilon_1-M_2-M_3'))\,\mathrm{E}_{\theta_0}[g(X_i;\theta_\infty,d)/f(X_i,\theta_0)]^{\alpha_1\alpha}\}^n\,. \end{aligned}$$

Similarly as in (4.31), the Chernoff's information inequality guarantees

$$\rho_4 = \inf_{0<\alpha<1} \{e^{\alpha\varepsilon_1} E_{\theta_0} [g(X_i;\theta_\infty, d)/f(X_i, \theta_0)]^{\alpha_1\alpha}\} < 1 .$$

That is,

$$(4.34) \qquad P_{\theta_0}\left\{ \sup_{t_1 \leqq t<\infty} \exp\left[\sum_{i=1}^{n} h_t(X_i; d) + nM_2 + nM_3'\right] > e^{-n\varepsilon_1}\right\} < \rho_4^n .$$

Therefore, from (4.22), (4.27), (4.32) and (4.34) we obtain the result of (4.21) for $n \geqq n_1$ and $\rho_1 < 1$ such that $n_1 \geqq -\varepsilon_1^{-1}\alpha_1 \log \eta$ and $\rho_1 = 3^{1/n_1}(\rho_2 \vee \rho_3 \vee \rho_4) < 1$.

THEOREM 4.3. *Under Assumptions A and C in Section 2, the PTMLE* $\{\hat{\theta}_{n:t};\ 0<t<\infty\}$ *is a.s. consistent for* θ_0*; this convergence is uniform in* $t:\ t_1 \leqq t < \infty$ *for every* $t_1 > 0$.

PROOF. Note that by Lemma 4.1 and 4.2, for every $d>0$ and η: $0<\eta<1$,

$$(4.35) \qquad P_{\theta_0}\{\sup [L_{n:t}(\theta)/L_{n:t}(\theta_0):\ t_1 \leqq t<\infty,\ \|\theta-\theta_0\|>d] > \eta\} \leqq \rho_0^n + \rho_1^n \qquad \text{for every } n \geqq n^* = n_0 \vee n_1 ,$$

where we take $M = M_1$. On the other hand, by (4.1), it holds that for every n and $t_1 > 0$

$$(4.36) \qquad \inf_{t_1 \leqq t<\infty} \{L_{n:t}(\hat{\theta}_{n:t})/L_{n:t}(\theta_0)\} \geqq 1 , \qquad \text{with probability } 1 .$$

From (4.35) and (4.36), we conclude that for every $d>0$ and $\eta>0$,

$$(4.37) \qquad P_{\theta_0}\left\{\sup_{t_1 \leqq t<\infty} \|\hat{\theta}_{n:t} - \theta_0\| \geqq d\right\} \leqq 2(\rho^*)^n , \qquad \text{for } n \geqq n^* ,$$

where $\rho^* = \rho_0 \vee \rho_1 < 1$. This completes the proof of the theorem.

5. Weak convergence of PTMLE

By virtue of Theorem 4.3 and the differentiability of the PT likelihood function in (3.1), it follows that the PTMLE, defined by (4.1), satisfies the PT likelihood equation (3.2), simultaneously in $t:\ t_1 \leqq t < \infty$, for every $t_1 > 0$ with the probability converging to 1 as $n \to \infty$. This implies that for any $d:\ 0<d \leqq d_0$ and any $t_1>0$, there exists a set $B_n \in \mathcal{B}_{n:\infty}$ with $P_{\theta_0}(B_n) = 1 - p_n$ (say), where

$$(5.1) \qquad p_n = P_{\theta_0}(B_n^c) \to 0 \qquad \text{as } n \to \infty ,$$

such that for every $x = (x_1, \cdots, x_n) \in B_n$

$$(5.2) \qquad \sup_{t_1 \leqq t<\infty} \|\hat{\theta}_{n:t}(x) - \theta_0\| < d$$

and

$$n^{-1/2}\Phi_{n:t}(\hat{\theta}_{n:t}(x))=0\,, \qquad \text{for any } t\in[t_1,\infty)\,. \tag{5.3}$$

We define

$$\dot{\phi}_t(x,\theta)=(\partial/\partial\theta')\phi_t(x,\theta)=(\partial^2/\partial\theta\partial\theta')\log f_t(x,\theta)\,; \tag{5.4}$$

$$\dot{\Phi}_{n:t}(\theta)=\sum_{i=1}^{n}\dot{\phi}_t(X_i,\theta)\,, \qquad t\in R^+\,. \tag{5.5}$$

Then, by the Taylor expansion, we have from (5.2) and (5.3),

$$0=n^{-1/2}\Phi_{n:t}(\hat{\theta}_{n:t})=n^{-1/2}\Phi_{n:t}(\theta_0)+n^{-1}\dot{\Phi}_{n:t}(\tilde{\theta}_{n:t})\{n^{1/2}(\hat{\theta}_{n:t}-\theta_0)\} \tag{5.6}$$

on B_n, where

$$\|\tilde{\theta}_{n:t}-\theta_0\|\leqq\|\hat{\theta}_{n:t}-\theta_0\|<d\,. \tag{5.7}$$

With these preliminaries, to study the weak convergence of the PTMLE, first, we consider the following.

LEMMA 5.1. *For every* η *and* ε: $0<\eta,\varepsilon<1$, *there exist positive numbers* d_2: $0<d_2<d_0$ *and* $n_0\geqq 1$ *such that for every* d: $0<d<d_2$ *and every* $n\geqq n_0$

$$P_{\theta_0}\{\sup[\|n^{-1}\dot{\Phi}_{n:t}(\theta)-n^{-1}\dot{\Phi}_{n:t}(\theta_0)\|:\ 0\leqq t<\infty,\ \|\theta-\theta_0\|<d]>\eta\}<\varepsilon\,. \tag{5.8}$$

PROOF.

$$\begin{aligned}(5.9)\quad &\|n^{-1}\dot{\Phi}_{n:t}(\theta)-\dot{\Phi}_{n:t}(\theta_0)\|\\ &\leqq n^{-1}\sum_{i=1}^{r_n(t)}\|\dot{\phi}(X_{n:i},\theta)-\dot{\phi}(X_{n:i},\theta_0)\|\\ &\quad+\bar{F}_n(t)\|(\partial/\partial\theta)\bar{\phi}(t,\theta)-(\partial/\partial\theta)\bar{\phi}(t,\theta_0)\|\\ &\leqq n^{-1}\sum_{i=1}^{n}u(X_i;\theta_0,d)\\ &\quad+(\bar{F}_n(t)/\bar{F}_{\theta_0}(t))\{\bar{F}_{\theta_0}(t)\|(\partial^2/\partial\theta\partial\theta')\log\bar{F}_\theta(t)-(\partial^2/\partial\theta\partial\theta')\log\bar{F}_{\theta_0}(t)\|\}\,,\end{aligned}$$

where the $u(x;\theta_0,d)$ are defined by (2.17) and $\bar{F}_n(t)$ is the empirical survival function. Thus, the left hand side of (5.8) is bounded from above by

$$\begin{aligned}(5.10)\quad &P_{\theta_0}\left\{n^{-1}\sum_{i=1}^{n}u(X_i;\theta_0,d)>\eta/2\right\}\\ &\quad+P_{\theta_0}\Big\{\sup[\bar{F}_n(t)/\bar{F}_{\theta_0}(t)\{\bar{F}_{\theta_0}(t)\|(\partial/\partial\theta)\bar{\phi}(t,\theta)-(\partial/\partial\theta)\bar{\phi}(t,\theta_0)\|\}:\\ &\qquad\qquad t\in R^+,\ \|\theta-\theta_0\|<d]>\eta/2\Big\}\,.\end{aligned}$$

By Assumption [B3], d can be chosen so small that

$$P_{\theta_0}\left\{n^{-1}\sum_{i=1}^{n}u(X_i;\theta_0,d)>\eta/2\right\}\leqq(\eta/2)^{-1}\,\mathrm{E}_{\theta_0}u(X_i;\theta_0,d)<\varepsilon/2\,. \tag{5.11}$$

Now, let us consider the value

$$(5.12)\qquad R(d)=\sup\{\bar{F}_{\theta_0}(t)\|(\partial^2/\partial\theta\partial\theta')\log\bar{F}_\theta(t)-(\partial^2/\partial\theta\partial\theta')\log\bar{F}_{\theta_0}(t)\|:\ \|\theta-\theta_0\|<d\}$$
$$\leqq\max\{R_T(d),\ \bar{R}_T(d)\}$$

where $\bar{R}_T(d)$ is defined in (2.19) and

$$(5.13)\qquad R_T(d)=\sup\{\bar{F}_{\theta_0}(t)\|(\partial^2/\partial\theta\partial\theta')\log\bar{F}_\theta(t)-(\partial^2/\partial\theta\partial\theta')\log\bar{F}_{\theta_0}(t)\|:\ \|\theta-\theta_0\|<d,\ 0\leqq t\leqq T\}\,.$$

Then, Assumption [B4] implies that T can be chosen so large that for any $M>0$ and all sufficiently small $d>0$

$$(5.14)\qquad \bar{R}_T(d)<\eta/(2M)\,.$$

From (2.41)-(2.42) and Lemma 3.2 we can choose d so small for any $T>0$ and $M>0$ that

$$(5.15)\qquad R_T(d)<\eta/(2M)\,.$$

Therefore, for any η and $M>0$ there exists $d_0>0$ such that for any $d\in(0,d_0)$

$$(5.16)\qquad R(d)<\eta/(2M)\,.$$

On the other hand, it follows by Chernoff and Rubin [3] that

(5.17) $\sup\{\bar{F}_n(t)/\bar{F}_{\theta_0}(t);\ t\in R^+\}$ is stochastically bounded.

Thus, from (5.16) and (5.17) we have that, for sufficiently large $M>0$ and sufficiently small $d>0$, the second term of (5.10) is bounded from above by

$$(5.18)\qquad P_{\theta_0}\left\{\sup_{t\in R^+}[(\bar{F}_n(t)/\bar{F}_{\theta_0}(t))R(d)]>\eta/2\right\}$$
$$\leqq P_{\theta_0}\left\{\sup_{t\in R^+}[(\bar{F}_n(t)/\bar{F}_{\theta_0}(t))]>M\right\}<\varepsilon/2\,.$$

From (5.10), (5.11) and (5.18) we conclude the proof of this lemma.

LEMMA 5.2. *For every* ε,η: $0<\varepsilon,\eta<1$, *there exists a positive integer* n_1 *such that for every* $n\geqq n_1$

$$(5.19)\qquad P_{\theta_0}\{\sup[\|n^{-1}\dot{\Phi}_{n:t}(\theta_0)+I_t(\theta_0)\|;\ t\in R^+]>\eta\}<\varepsilon\,.$$

PROOF. By (2.4) and (2.41), it holds that

$$(5.20)\qquad \|n^{-1}\dot{\Phi}_{n:t}(\theta_0)+I_t(\theta_0)\|$$

$$=\left\| n^{-1}\sum_{i=1}^{r_n(t)} \dot\phi(X_{n:i},\theta_0)+\bar F_n(t)\int_t^\infty \dot\phi(x,\theta_0)F_{\theta_0}(dx)/\bar F_{\theta_0}(t)+I(\theta_0) + (\bar F_n(t)/\bar F_{\theta_0}(t)-1)\bar I_t(\theta_0)\right\|$$

$$\leqq \left\| n^{-1}\sum_{i=1}^{n}\dot\phi_t(X_i,\theta_0)+I(\theta_0)\right\| + |\bar F_n(t)/\bar F_{\theta_0}(t)-1|\,\|\bar I_t(\theta_0)\|\,,$$

where $\dot\phi_t(x,\theta_0)$ is the projection of $\dot\phi(x,\theta_0)$; that is,

$$\dot\phi_t(x,\theta_0)=\mathrm{E}_{\theta_0}[\dot\phi(X_1,\theta_0)|\mathcal{B}_{1:t}] = \begin{cases} \dot\phi(x,\theta_0)\,, & \text{if } 0\leqq x\leqq t\,,\\ \int_t^\infty \dot\phi(x,\theta_0)F_{\theta_0}(dx)/\bar F_{\theta_0}(t)\,, & \text{if } t<x<\infty\,. \end{cases} \tag{5.21}$$

Therefore, the left hand side of (5.19) is bounded from above by

$$P_{\theta_0}\left\{\sup\left[\left\| n^{-1}\sum_{i=1}^{n}\dot\phi_t(X_i,\theta_0)+I(\theta_0)\right\|;\ t\in R^+\right]>\eta/2\right\} + P_{\theta_0}\{\sup[|\bar F_n(t)/\bar F_{\theta_0}(t)-1|\,\|\bar I_t(\theta_0)\|;\ t\in R^+]>\eta/2\}\,. \tag{5.22}$$

Since $\left\{n^{-1}\sum_{i=1}^{n}\dot\phi_t(X_i,\theta_0)+I(\theta_0),\ \mathcal{B}_{n:t};\ t\in R^+\right\}$ is a zero mean martingale for every n, we have by the martingale inequality that

$$\begin{aligned} &P_{\theta_0}\left\{\sup\left[\left\| n^{-1}\sum_{i=1}^{n}\dot\phi_t(X_i,\theta_0)+I(\theta_0)\right\|;\ t\in R^+\right]>\eta/2\right\}\\ &\leqq 2\eta^{-1}\,\mathrm{E}_{\theta_0}\left\| n^{-1}\sum_{i=1}^{n}\dot\phi(X_i,\theta_0)+I(\theta_0)\right\|^2\\ &=2\eta^{-1}n^{-1}\,\mathrm{E}_{\theta_0}\|\dot\phi(X_i,\theta_0)+I(\theta_0)\|^2<\varepsilon/2 \end{aligned} \tag{5.23}$$

for all sufficiently large n.

Next, by (2.39) and (2.40), for any $\eta>0$ and $M>0$ we can choose T enough large to let

$$\|\bar I_t(\theta_0)\|<\eta/(2M)\,,\qquad \text{for every } t>T\,. \tag{5.24}$$

By the result due to Chernoff and Rubin [3] that $\sup[|\bar F_n(t)/\bar F_{\theta_0}(t)-1|;\ t\in R^+]$ is stochastically bounded, this implies that

$$\begin{aligned} &P_{\theta_0}\{\sup[|\bar F_n(t)/\bar F_{\theta_0}(t)-1|\,\|\bar I_t(\theta_0)\|;\ T<t<\infty]>\eta/2\}\\ &\leqq P_{\theta_0}\{\sup[|\bar F_n(t)/\bar F_{\theta_0}(t)-1|;\ t\in R^+]>M\}<\varepsilon/4\,. \end{aligned} \tag{5.25}$$

Furthermore, note that $\{|\bar F_n(t)/\bar F_{\theta_0}(t)-1|,\ \mathcal{B}_{n:t};\ t\in R^+\}$ is a submartingale for each n. Then, it is shown by the submartingale inequality that

$$\begin{aligned} &P_{\theta_0}\{\sup[|\bar F_n(t)/\bar F_{\theta_0}(t)-1|\,\|\bar I_t(\theta_0)\|;\ 0\leqq t\leqq T]>\eta/2\}\\ &\leqq P_{\theta_0}\{\sup[|\bar F_n(t)/\bar F_{\theta_0}(t)-1|;\ 0\leqq t\leqq T]>\eta/(2\|I(\theta_0)\|) \end{aligned} \tag{5.26}$$

$$\leqq (2\|I(\theta_0)\|/\eta)^2 \mathrm{E}_{\theta_0} |\bar{F}_n(T)/\bar{F}_{\theta_0}(T)-1|^2$$
$$=(2\|I(\theta_0)\|/\eta)^2 n^{-1}(F_{\theta_0}(T)/\bar{F}_{\theta_0}(T)) < \varepsilon/4$$

for all sufficiently large n. (5.25) and (5.26) lead to

$$P_{\theta_0}\{\sup [|\bar{F}_n(t)/\bar{F}_{\theta_0}(t)-1|\|\bar{I}_t(\theta_0)\|;\ t \in R^+] > \eta/2\} < \varepsilon/2 . \tag{5.27}$$

This together with (5.22) and (5.23) completes the proof of this lemma.

THEOREM 5.3. *Under the regularity conditions of Section 2, for any* $t_1>0$, *the PTMLE* $\{n^{1/2}(\hat{\theta}_{n:t}-\theta_0),\ t_1 \leqq t < \infty\}$ *converges weakly in* $D^k[t_1, \infty)$ *to the Gaussian vector process*

$$[I_t(\theta_0)]^{-1}\Phi_t^\circ(\theta_0) = [I_t(\theta_0)]^{-1}\int_0^\infty \phi_t(x, \theta_0) W^\circ(F_{\theta_0}(dx)) , \qquad t_1 \leqq t < \infty , \tag{5.28}$$

where $\Phi_t^\circ(\theta_0)$ *is defined by* (3.10)–(3.11).

PROOF. By virtue of (5.1)–(5.7), we have (on the set B_n) with probability arbitrarily closed to 1 (for large n),

$$-n^{-1}\Phi_{n:t}(\tilde{\theta}_{n:t})\{n^{1/2}(\hat{\theta}_{n:t}-\theta_0)\} = n^{-1/2}\Phi_{n:t}(\theta_0) , \qquad t_1 \leqq t < \infty . \tag{5.29}$$

Since by Theorem 3.1,

$$\sup \{\|n^{-1/2}\Phi_{n:t}(\theta_0)\|;\ t_1 \leqq t < \infty\} = 0_p(1) , \tag{5.30}$$

it holds by (5.29) that for any $t_1>0$

$$\sup \{\|n^{-1}\dot{\Phi}_{n:t}(\tilde{\theta}_{n:t})(n^{1/2}(\hat{\theta}_{n:t}-\theta_0))\|;\ t_1 \leqq t < \infty\} = 0_p(1) . \tag{5.31}$$

Hence, by using Lemma 5.1, 5.2 and (5.31), we conclude that

$$\sup \{\|I_t(\theta_0)(n^{1/2}(\hat{\theta}_{n:t}-\theta_0))\|;\ t_1 \leqq t < \infty\} = 0_p(1) , \tag{5.32}$$

which, in turn, along with (5.29), ensures that

$$\sup \{\|n^{-1/2}\Phi_{n:t}(\theta_0) - I_t(\theta_0)(n^{1/2}(\hat{\theta}_{n.t}-\theta_0))\|;\ t_1 \leqq t < \infty\} \to 0 , \tag{5.33}$$

in probability, as $n \to \infty$. Since $\|I_{t_1}(\theta_0)\| \leqq \|I_t(\theta_0)\| \leqq \|I(\theta_0)\| < \infty$ for every $t_1 \leqq t < \infty$, we have, by (5.33),

$$\sup \{\|n^{1/2}(\hat{\theta}_{n:t}-\theta_0) - I_t^{-1}(\theta_0)(n^{-1/2}\Phi_{n:t}(\theta_0))\|;\ t_1 \leqq t < \infty\} \to 0 \tag{5.34}$$

in probability as $n \to \infty$. Therefore, (5.28) follows from (5.22) and Theorem 3.1.

6. Discussions and remarks

Under the regularity conditions mentioned in Section 2, we have the result (same as in Sen and Tsong [11]) that $\{n^{-1}\Phi_{n:t}(\theta_0)-I_t(\theta_0)b\}$ converges weakly in $D^k[t_1, \infty)$ to $\Phi_t^\circ(\theta_0)$ as $n \to \infty$ when alternative $\{H_n: \theta_n=\theta_0+n^{-1}b,\ b \in R^k\}$ holds. By virtue of contiguity of $\{P_{\theta_0}\}$ and $\{P_{\theta_n}\}$, this and (5.34) imply that $\{n^{1/2}(\hat{\theta}_{n:t}-\theta_n);\ t_1 \leqq t < \infty\}$ convergences weakly in $D^k[t_1, \infty)$ to $I_t^{-1}(\theta_0)\Phi_t^\circ(\theta_0)$ when $\{H_n\}$ holds. Note that the limit distribution is independent of local alternative $b \in R^k$. In other words, the PTMLE is regular in $D^k[t_1, \infty)$.

It is worth remarking that several distribution functions appeared in the survival analysis (for example, normal, log normal, inverse Gaussian, Gamma distributions and so on,) satisfy the regularity conditions mentioned in Section 2. Assumptions C which depend on the parametrization of the distribution enable us to compactify the parameter space. Therefore, we have to check that the regularity conditions hold only in any compact subset of the parameter space. It is easy to see that Assumptions A and B, except for [B4], hold. However, we have to prove [B4] case by case.

OSAKA UNIVERSITY
UNIVERSITY OF NORTH CAROLINA AT CHAPEL HILL

REFERENCES

[1] Bangdiwala, S. I. (1983). A progressively censored sequential likelihood ratio test for general parametric hypothesis, *Commun. Statist.-Sequential Analysis*, **2** (1), 1-25.

[2] Chernoff, H. (1952). A measure of asymptotic efficiency for test of a hypothesis based on the sum of observations, *Ann. Math. Statist.*, **23**, 493-507.

[3] Chernoff, H. and Rubin, H. (1956). The estimation of the location of a discontinuity in density, *Proc. of The Third Berkeley Symp. Math. Statist. Prob.* (ed. Neyman, J.), **1**, 19-37.

[4] Gardiner, J. C. (1981). Convergence of progressively censored likelihood ratio processes in life-testing, *Sankhya*, **43**, A, 37-51.

[5] Inagaki, N. (1973). Asymptotic relations between the likelihood estimating function and the maximum likelihood estimator, *Ann. Inst. Statist. Math.*, **25**, 1-26.

[6] Inagaki, N. (1983). The decomposition of the Fisher information, *Ann. Inst. Statist. Math.*, **35**, 151-165.

[7] Karlin, S. and Taylor, H. M. (1981). *A First Course in Stochastic Process*, second edition, Academic Press.

[8] Rao, M. M. (1981). *Foundations of Stochastic Analysis*, Academic Press.

[9] Sen, P. K. (1976). Weak convergence of progressively censored likelihood ratio statistics and its role in asymptotic theory of life testing, *Ann. Statist.*, **4**, 1247-1257.

[10] Sen, P. K. (1981). *Sequential Nonparametrics*, John Wiley and Sons, New York.

[11] Sen, P. K. and Tsong, Y. (1981). An invariance principle for progressively truncated likelihood ratio statistics, *Metrika*, **28**, 165-177.