Ann. Inst. Statist. Math.
29 (1977), Part A, 165-187

THE WEAK CONVERGENCE OF THE LIKELIHOOD RATIO RANDOM FIELDS FOR MARKOV OBSERVATIONS

YOSIHIKO OGATA AND NOBUO INAGAKI

(Received Mar. 24, 1976; revised Oct. 15, 1976)

1. Introduction

Let $(\Omega, \mathfrak{A})$ be a measurable space and Θ be a subset of a k-dimensional Euclidean space. For each $\theta \in \Theta$ let P_θ be a probability measure on $(\Omega, \mathfrak{A})$. We assume that for every $\theta \in \Theta$ $\{X_n: n \geqq 0\}$ is a stationary Markov process defined on $(\Omega, \mathfrak{A}, P_\theta)$ into $(\mathcal{X}, \mathfrak{B})$ where $(\mathcal{X}, \mathfrak{B})$ denotes some-dimensional Euclidean space with the Borel σ-algebra. For each $\theta \in \Theta$, $x \in \mathcal{X}$ and $B \in \mathfrak{B}$ let $Q_\theta(x, B)$ be a transition probability of the stationary Markov process $\{X_n: n \geqq 0\}$. Suppose that for each $\theta \in \Theta$ there exists a uniquely determined invariant probability measure $q_\theta(\cdot)$ on $\mathfrak{B}$. Without loss of generality we assume that $(\Omega, \mathfrak{A})$ is the infinite Cartesian product $\prod_{i=1}^{\infty} (\mathcal{X}, \mathfrak{B})$, and that P_θ is the probability measure induced in $\mathfrak{A}$ by $q_\theta(\cdot)$ and $Q_\theta(\cdot, \cdot)$. Suppose further that there is a σ-finite measure μ on $\mathfrak{B}$ such that $q_\theta(\cdot)$ and $Q_\theta(x, \cdot)$ for each $x \in \mathcal{X}$ are absolutely continuous with respect to μ. That is,

$$(1.1) \qquad q_\theta(B) = \int_B f(y; \theta)\mu(dy) , \qquad Q_\theta(x, B) = \int_B f(x, y; \theta)\mu(dy)$$

hold for any $B \in \mathfrak{B}$ and $x \in \mathcal{X}$.

Denote the likelihood ratio statistic by

$$(1.2) \qquad Z_n(h) = Z_n(h, \omega) = \prod_{i=1}^{n} \{f(X_{i-1}(\omega), X_i(\omega); \theta_0 + h/\sqrt{n})/ f(X_{i-1}(\omega), X_i(\omega); \theta_0)\}$$

for $\omega \in \Omega$ and θ_0, $\theta_0 + h/\sqrt{n} \in \Theta$, where θ_0 is the true parameter (which is any one of Θ but fixed). We shall regard $h \rightleftarrows Z_n(h)$ as a random field with multi-dimensional parameter h, $\theta_0 + h/\sqrt{n} \in \Theta$.

In this paper we shall study asymptotic behaviors of the likelihood ratio statistic for Markov processes from the view of the weak convergence of the random field. LeCam [9] and Ibragimov-Khas'minskii [6] investigate those in the case of one-dimensional parameter and i.i.d.

observations under assumptions essentially different from those in the present paper. The authors of the present paper [7] discuss asymptotic behaviors of the likelihood ratio random field and functionals on it in the case of multi-dimensional parameter but i.i.d. observations.

In Section 2 notations, assumptions and some remarks are stated. In Section 3 we collect the main results in the present paper and conclude the weak convergence of likelihood ratio random fields to some degenerated Gaussian random field. In Section 4 we give those proofs and several lemmas. Finally in Section 5 we discuss a certain statistical Markovian model and mixing condition of Assumption (A6) (below). Several examples are given which are valuable for the analysis of time series. See [10] for the discussions about the same applications as in Inagaki and Ogata [7].

2. Assumptions and some remarks

In this section we state three groups of assumptions and give some remarks. Assumptions A are primitive, B are local at the true parameter, and C are grobal with respect to θ. Denote the true parameter by θ_0 which is supposed to be any interior point of Θ but fixed. Let $|\cdot|$ be the maximum norm, i.e. for $\theta^{(i)} \in R^1$, $|\theta^{(i)}|=$the absolute value of $\theta^{(i)}$ and for $\theta=(\theta^{(1)},\cdots,\theta^{(k)})^T$, $|\theta|=\max\{|\theta^{(1)}|,\cdots,|\theta^{(k)}|\}$.

ASSUMPTIONS A.

(A1) The parameter space Θ is a subset of R^k.

(A2) For each $\theta\in\Theta$, Q_θ has a derivative with respect to a σ-finite measure μ:

$$f(x, y; \theta)=Q_\theta(x, dy)/\mu(dy) .$$

(A3) For each $\theta\in\Theta$, $f(x, y; \theta)$ has a common support.

(A4) $f(x, y; \theta)$ is $\mathfrak{B}\times\mathfrak{B}$-measurable for every $\theta\in\Theta$ and continuous in Θ for a.s. $[\mu\times\mu](x, y)$.

(A5) For any $\theta_1, \theta_2\in\Theta$, $\theta_1\neq\theta_2$ implies

$$\int |f(x, y; \theta_1)-f(x, y; \theta_2)|\mu(dy)>0$$

for a.s. $[\mu]x$.

Denote the n-step transition probability and its derivative with respect to μ by $Q_\theta^{(n)}(x, B)$ and $f^{(n)}(x, y; \theta)$, respectively:

$$(2.1)\quad \begin{aligned} Q_\theta^{(n)}(x, B)&=\int_B f^{(n)}(x, y; \theta)\mu(dy) ,\\ f^{(n)}(x, y; \theta)&=\overbrace{\int\cdots\int}^{n-1} f(x, x_1; \theta)\cdots f(x_{n-1}, y; \theta)\mu(dx_1)\cdots\mu(dx_{n-1}) \end{aligned}$$

for $x, y \in \mathcal{X}$, $B \in \mathfrak{B}$ and $n \geqq 1$. Set

(2.2) $$\phi(n)=\iint f(x, \theta_0)\,|f^{(n)}(x, y; \theta_0)-f(y; \theta_0)|\,\mu(dx)\mu(dy) .$$

(A6) There exists a positive constant $p>1$ such that

$$\sum_{n=1}^{\infty} \phi(n)^{1/p}<\infty .$$

There is a neighborhood of θ_0,

$$U_0=U_{d_0}(\theta_0)=\{\theta;\,|\theta-\theta_0|<d_0\}\ \text{(say)} \subset \Theta$$

satisfying the following conditions.

ASSUMPTIONS B.

(B1) $\log f(x, y; \theta)$ is continuously differentiable in U_0 for a.s. $[\mu\times\mu]$ (x, y): set

(2.3) $$\eta(x, y; \theta)=\frac{\partial}{\partial\theta} \log f(x, y; \theta)=\left(\frac{\partial}{\partial\theta^{(1)}}, \cdots, \frac{\partial}{\partial\theta^{(k)}}\right)^t \log f(x, y; \theta) .$$

(B2) For each $\theta \in U_0$, $\eta(x, y; \theta)$ is $\mathfrak{B}\times\mathfrak{B}$-measurable.

(B3) For each $\theta \in U_0$,

(2.4) $$\lambda(\theta)=\mathrm{E}_{\theta_0}\, \eta(X_0, X_1; \theta)\ \text{(say)} , \qquad \text{exists} .$$

(B4) For every $\theta \in U_0$, the matrix

(2.5) $$\Gamma(\theta)=\mathrm{E}_\theta\, \{\eta(X_0, X_1; \theta)\eta(X_0, X_1; \theta)^t\}\ \text{(say)}$$

exists and is continuous at θ_0. $\Gamma(\theta_0)$ is positive definite. ($\Gamma(\theta)$ is so called Fisher's Information matrix.)

(B5) $\lambda(\theta)$ is continuously differentiable at θ_0:

(2.6) $$\Lambda(\theta_0)=\frac{\partial}{\partial\theta}\lambda(\theta_0)=\left(\frac{\partial\lambda^{(i)}(\theta_0)}{\partial\theta^{(j)}}\right)_{i, j=1, \cdots, k}\ \text{(say)} ,$$
$$-\Lambda(\theta_0)=\Gamma(\theta_0) .$$

Set

(2.7) $$u(x, y; \theta, d)=\sup\,\{|\eta(x, y; \tau)-\eta(x, y; \theta)|;\ |\tau-\theta|<d\}$$

for every $\theta \in U_0$ and let q be such that

(2.8) $$1/p+1/q=1$$

for $p>1$ in Assumption (A6).

(B6) There are positive constants H_s, $s=1$, 2 and $2q$, such that

$$E_{\theta_0} u(X_0, X_1; \theta, d)^s < H_s d$$

for $|\theta-\theta_0|+d<d_0$, $d>0$.

(B7) $$E_{\theta_0}(|\eta(X_0, X_1; \theta_0)|^{2q}) = H_0 \text{ (say)} < \infty .$$

ASSUMPTIONS C.

Let $(\bar{\Theta}, \delta)$ be a metric space with the metric

(2.9) $$\delta(\theta_1, \theta_2) = |\theta_1-\theta_2| / \{1+|\theta_1-\theta_2|\}$$

for $\theta_1, \theta_2 \in \bar{\Theta}$, satisfying the followings.

(C1) $(\bar{\Theta}, \delta)$ is the Bahadur compactification of Θ (see Bahadur [1], p. 21) such that

(i) $\bar{\Theta}$ is compact.

(ii) $\Theta \subset \bar{\Theta}$ and Θ is everywhere dense in $\bar{\Theta}$.

Put

(2.10) $$\begin{aligned} g(x, y; \bar{\theta}, d) &= \sup\{f(x, y; \theta):\ \theta \in \Theta,\ \delta(\theta, \bar{\theta}) < d\} , \\ g(x, y; \theta_\infty, d) &= \sup\{f(x, y; \theta):\ \theta \in \Theta,\ \delta(\theta_0, \theta) > 1-d\} \end{aligned}$$

for $\bar{\theta} \in \bar{\Theta}$ with $\delta(\theta_0, \bar{\theta}) < 1$ and $\theta_\infty \in \bar{\Theta}$ with $\delta(\theta_0, \theta_\infty) = 1$, respectively.

(iii) For each $\bar{\theta} \in \bar{\Theta}$, there exists $d_1 = d_1(\bar{\theta}) > 0$ such that for each d, $0 \leqq d \leqq d_1$, $g(x, y; \bar{\theta}, d)$ is $\mathfrak{B} \times \mathfrak{B}$-measurable, $0 \leqq g \leqq \infty$.

(iv) For each $\bar{\theta} \in \bar{\Theta}$ and a.s. $[\mu]x$,

$$\int g(x, y; \bar{\theta}, 0) \mu(dy) \leqq 1$$

where

(2.11) $$g(x, y; \bar{\theta}, 0) = \lim_{d \to 0} g(x, y; \bar{\theta}, d) .$$

(C2) If $\bar{\theta} \neq \theta_0$, then

$$\int |g(x, y; \bar{\theta}, 0) - f(x, y; \theta_0)| \mu(dy) > 0$$

for a.s. $[\mu]x$.

(C3) For every $\bar{\theta} \in \bar{\Theta}$, there exists $d = d(\bar{\theta})$, $0 < d \leqq d_1$, such that

$$E_{\theta_0}[|\log\{f(X_0, X_1; \theta_0)/g(X_0, X_1; \bar{\theta}, d)\}|^{2q}] < \infty$$

where q is such as in (2.9).

(C4) For $\theta_\infty \in \bar{\Theta}$ with $\delta(\theta_\infty, \theta_0) = 1$,

$$\lim_{d \to 0} d^2 E_{\theta_0} \log\{f(X_0, X_1; \theta_0)/g(X_0, X_1; \theta_\infty, d)\} > 0 ,$$

$$\varlimsup_{d\to\infty} d^2[\mathrm{E}\,|\log\{f(X_0, X_1; \theta_0)/g(X_0, X_1; \theta_\infty, d)\} \\ -\mathrm{E}_{\theta_0}\log\{f(X_0, X_1; \theta_0)/g(X_0, X_1; \theta_\infty, d)\}|^{2q}]^{1/q}<\infty$$

From now on let us agree to write

$$\mathrm{E}(\cdot)\,,\quad \mathrm{P}(\cdot)\,,\quad q(\cdot) \text{ and } Q(\cdot\,,\cdot)$$

simply, instead of

$$\mathrm{E}_{\theta_0}(\cdot)\,,\quad \mathrm{P}_{\theta_0}(\cdot)\,,\quad q_{\theta_0}(\cdot) \text{ and } Q_{\theta_0}(\cdot\,,\cdot)\,,$$

respectively, unless especially mentioned.

Remarks.

(a) Set

$$Y(h)=Y(X_0, X_1; h)=f(X_0, X_1; \theta_0+h)^{1/2}/f(X_0, X_1; \theta_0)^{1/2}-1\,.$$

Then it is proved in the same way as in Lemma 2.1 of Inagaki and Ogata [7] that for a.s.

$$(2.12)\qquad \frac{1}{\varepsilon}Y(\varepsilon h)\to\frac{1}{2}h^t\eta(X_0, X_1; \theta_0)\,,\qquad \text{in probability}$$

and

$$(2.13)\qquad \lim_{\varepsilon\to0}\mathrm{E}\left[\left\{\frac{1}{\varepsilon}Y(\varepsilon h)\right\}^2\Big|X_0\right]=\mathrm{E}\left[\left\{\frac{1}{2}h^t\eta(X_0, X_1; \theta_0)\right\}^2\Big|X_0\right]\,.$$

Thus we have the same results as in LeCam [9], p. 807 that $Y(h)$ is differentiable in quadratic mean at 0 for a.s.:

$$(2.14)\qquad \mathrm{E}\left[\left\{\frac{1}{\varepsilon}Y(\varepsilon h)-\frac{1}{2}h^t\eta(X_0, X_1; \theta_0)\right\}^2\Big|X_0\right]\to0\,,\qquad \text{as } \varepsilon\to0\,,$$

and hence that

$$(2.15)\qquad \begin{aligned}&\mathrm{E}\,[\eta(X_0, X_1; \theta_0)\,|\,X_0]=0 \qquad \text{for a.s.}\,,\\ &\lambda(\theta_0)=\mathrm{E}\,\eta(X_0, X_1; \theta_0)=0\,.\end{aligned}$$

(b) From Assumption (B6), for $|\tau-\theta|<d$ and $|\theta-\theta_0|+d\leqq d_0$

$$|\lambda(\tau)-\lambda(\theta)|\leqq\mathrm{E}\,u(X_0, X_1; \theta, d)\leqq H_1 d\,,$$

that is, $\lambda(\theta)$ is uniformly continuous on U_0.

(c) Since $\phi(n)$ is the expectation of the conditional total variation measure of the signed measure $Q^{(n)}(x, dy)-q(dy)$, we can easily see that the stationary Markov process $\{X_n\}$ is strongly mixing with the mixing coefficient $\phi(n)$ and hence ergodic. (See Proposition 1 in Davydov [4], for example.) Denote the Kullback-Leibler information by

(2.16) $$K(\theta)=-\mathrm{E}\log\{f(X_0, X_1; \theta)/f(X_0, X_1; \theta_0)\}, \qquad \theta\in\Theta$$

and let

(2.17) $$\bar{K}(\theta)=-\mathrm{E}\log\{g(X_0, X_1; \theta_0, 0)/f(X_0, X_1; \theta_0)\}, \qquad \theta\in\bar{\Theta}.$$

(d) From Assumption (A4) and the definition (2.11) of $g(x, y; \theta, 0)$, we see that

$$g(x, y; \theta, 0)=f(x, y; \theta) \qquad \text{for } \theta\in\Theta$$

and so that $g(x, y; \theta, 0)$ is an extension of $f(x, y; \theta)$ on $\mathcal{X}\times\mathcal{X}\times\Theta$ to a function on $\mathcal{X}\times\mathcal{X}\times\bar{\Theta}$. Thus, $\bar{K}(\theta)$ on $\bar{\Theta}$ is regarded as an extension of $K(\theta)$ on Θ.

(e) From Assumption (C1)-(iv) and (C2), it follows that

$$0<\bar{K}(\theta)\leqq\infty, \qquad \text{for } \theta\,(\neq\theta_0)\in\bar{\Theta}.$$

(f) From Assumption (C3) and Lebesgue's dominated convergence theorem, it follows that

$$\lim_{d\to 0}\mathrm{E}\log\{g(X_0, X_1; \theta, d)/f(X_0, X_1; \theta_0)\}=-\bar{K}(\theta), \qquad \text{for } \theta\in\bar{\Theta},$$

and hence from Remark (e) that for $\theta\in\bar{\Theta}$ there is $d=d(\theta)$, $0<d<d_1$ satisfying

(2.18) $$-\infty<\mathrm{E}\log\{g(X_0, X_1; \theta, d)/f(X_0, X_1; \theta_0)\}<-\frac{1}{2}\bar{K}(\theta)<0.$$

The following lemmas are fundamental in this paper.

LEMMA 2.1. *Let* $\alpha(x, y)$ *and* $\beta(x, y)$ *be* $\mathfrak{B}\times\mathfrak{B}$*-measurable real valued functions such that* $\mathrm{E}\{|\alpha(X_0, X_1)|^{q_1}\}<\infty$ *and* $\mathrm{E}\{|\beta(X_0, X_1)|^{r_1}\}<\infty$ *for some* $p_1, q_1, r_1>1$ *with*

(2.19) $$1/p_1+1/q_1+1/r_1=1.$$

Then, for $i\geqq 1$

(2.20) $$\begin{aligned}|\mathrm{E}\{\alpha(X_0, X_1)\beta(X_i, X_{i+1})\}-\mathrm{E}\{\alpha(X_0, X_1)\}\,\mathrm{E}\{\beta(X_0, X_1)\}|\\ \leqq 2\phi(i-1)^{1/p_1}\{\mathrm{E}|\alpha(X_0, X_1)|^{q_1}\}^{1/q_1}\{\mathrm{E}|\beta(X_0, X_1)|^{r_1}\}^{1/r_1},\end{aligned}$$

where $\phi(k)$, $k\geqq 1$ *are those defined in* (2.2) *and*

(2.21) $$\phi(0)=1 \qquad \text{(say)}.$$

PROOF. Since $|\mathrm{E}\{\alpha(X_0, X_1)\beta(X_i, X_{i+1})\}|$ and $|\mathrm{E}\{\alpha(X_0, X_1)\}\,\mathrm{E}\{\beta(X_0, X_1)\}|$ are dominated by $\{\mathrm{E}|\alpha(X_0, X_1)|^{q_1}\}^{1/q_1}\{\mathrm{E}|\beta(X_0, X_1)|^{r_1}\}^{1/r_1}$, the inequality (2.20) holds if $\phi(i-1)^{1/p_1}=1$. Thus, it is sufficient to show that the inequality holds for $i\geqq 2$.

Consider the total variation measure $V_i(x, B)$ of the signed measure $Q^{(i)}(x, B)-q(B)$. Then,

$$V_i(x, B)=\int_B |f^{(i)}(x, y; \theta_0)-f(y; \theta_0)|\mu(dy) \tag{2.22}$$

and

$$V_i(x, B)\leqq Q^{(i)}(x, B)+q(B) \tag{2.23}$$

for every $x \in \mathcal{X}$ and $B \in \mathfrak{B}$. By Markov property and (2.22), it holds that

$$\begin{aligned} I_i &= |\mathrm{E}\{\alpha(X_0, X_1)\beta(X_i, X_{i+1})\} - \mathrm{E}\{\alpha(X_0, X_1)\}\mathrm{E}\{\beta(X_0, X_1)\}| \\ &= \left|\int \alpha(x, y)\beta(u, v)q(dx)Q(x, dy)\{Q^{(i-1)}(y, du)-q(du)\}Q(u, dv)\right| \\ &\leqq \int |\alpha(x, y)\beta(u, v)|q(dx)Q(x, dy)V_{i-1}(y, du)Q(u, dv) . \end{aligned} \tag{2.24}$$

Set

$$W_i(dx, dy, du, dv)=q(dx)Q(x, dy)V_i(y, du)Q(u, dv) . \tag{2.25}$$

Then, it follows from (2.23) that

$$dW_i\leqq q(dx)Q(x, dy)\{Q^{(i)}(y, du)+q(du)\}Q(u, dv) . \tag{2.26}$$

Applying Holder's inequality in (2.24), we have that

$$I_k\leqq\left\{\int dW_{k-1}\right\}^{1/p_1}\left\{\int |\alpha(x, y)|^{q_1}dW_{k-1}\right\}^{1/q_1}\{|\beta(u, v)|^{r_1}dW_{k-1}\}^{1/r_1}$$

and from the definitions $\phi(i)$ and (2.26) that

$$\begin{aligned} I_k&\leqq\phi(i-1)^{1/p_1}\{2\,\mathrm{E}|\alpha(X_0, X_1)|^{q_1}\}^{1/q_1}\{2\,\mathrm{E}|\beta(X_0, X_1)|^{r_1}\}^{1/r_1} \\ &\leqq 2\phi(i-1)^{1/p_1}\{\mathrm{E}|\alpha(X_0, X_1)|^{q_1}\}^{1/q_1}\{\mathrm{E}|\beta(X_0, X_1)|^{r_1}\}^{1/r_1} . \end{aligned}$$

Hence, the proof is complete.

LEMMA 2.2. *Let $\alpha(x, y)$ be a k-dimensional vector valued and $\mathfrak{B}\times\mathfrak{B}$-measurable function such that* $\mathrm{E}\{|\alpha(X_0, X_1)|^{2q}\}<\infty$ *for p and q in Assumption* (A6) *and* (2.8). *Then,*

$$\begin{aligned} &\mathrm{E}\left[\left|\sum_{i=1}^{n}\{\alpha(X_{i-1}, X_i)-\mathrm{E}\,\alpha(X_0, X_1)\}\right|^2\right] \\ &\qquad\leqq nk\{\mathrm{E}|\alpha(X_0, X_1)|^{2q}\}^{1/q}\left\{1+4\sum_{i=1}^{n-1}\phi(i-1)^{1/p}\right\} . \end{aligned}$$

PROOF. Put $\alpha(x, y)=(\alpha^{(1)}(x, y),\cdots, \alpha^{(k)}(x, y))$, then by the stationary property, we have

$$\mathrm{E}\left[\left|\sum_{i=1}^{n}\{\alpha(X_{i-1}, X_i)-\mathrm{E}\,\alpha(X_0, X_1)\}\right|^2\right] \tag{2.27}$$

$$\leqq \sum_{r=1}^{k} \mathrm{E}\left[\sum_{i=1}^{n} \{\alpha^{(r)}(X_{i-1}, X_i) - \mathrm{E}\, \alpha^{(r)}(X_0, X_1)\}\right]^2$$
$$= \sum_{r=1}^{k} \Big[n\, \mathrm{E}\, \{\alpha^{(r)}(X_0, X_1)\}^2 - n(\mathrm{E}\, \alpha^{(r)}(X_0, X_1))^2$$
$$+ 2 \sum_{i=1}^{n-1} (n-i)\{\mathrm{E}\, \alpha^{(r)}(X_0, X_1)\alpha^{(r)}(X_i, X_{i+1}) - (\mathrm{E}\, \alpha(X_0, X_1))^2\}\Big] .$$

Taking $p_1 = p$, $q_1 = r_1 = 2q$ and $\beta(x, y) = \alpha(x, y) = \alpha^{(r)}(x, y)$ in Lemma 2.1, we have

$$(2.28) \qquad \begin{aligned} &|\mathrm{E}\, \alpha^{(r)}(X_0, X_1)\alpha^{(r)}(X_i, X_{i+1}) - (\mathrm{E}\, \alpha^{(r)}(X_0, X_1))^2| \\ &\leqq 2\phi(i-1)\{\mathrm{E}\,|\alpha^{(r)}(X_0, X_1)|^{2q}\}^{1/q} \\ &\leqq 2\phi(i-1)\{\mathrm{E}\,|\alpha(X_0, X_1)|^{2q}\}^{1/q} \end{aligned}$$

for $1 \leqq r \leqq k$ and $i \geqq 1$. Thus, we conclude from (2.27) and (2.28) that

$$\begin{aligned} &\mathrm{E}\left[\left|\sum_{i=1}^{n} \{\alpha(X_{i-1}, X_i) - \mathrm{E}\, \alpha(X_0, X_1)\}\right|^2\right] \\ &\leqq k\left[n\, \mathrm{E}\,|\alpha(X_0, X_1)|^2 + 4\{\mathrm{E}\,|\alpha(X_0, X_1)|^{2q}\}^{1/q} \sum_{i=1}^{n-1} (n-i)\phi(i-1)^{1/p}\right] \\ &\leqq nk\{\mathrm{E}\,|\alpha(X_0, X_1)|^{2q}\}^{1/q}\left\{1 + 4\sum_{i=1}^{n-1} \phi(i-1)^{1/p}\right\} . \end{aligned}$$

The proof of this lemma is complete.

Set

$$(2.29) \qquad \xi_n(\theta) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \eta(X_{i-1}, X_i ; \theta) .$$

From (2.15) of Remark (a) and Markov property, we have that $\xi_n(\theta_0)$, $n = 1, 2, \cdots$ form a martingale and hence, obtain the following theorem by Billingsley's Theorem ([2], p. 788) and Assumption (B4).

THEOREM 2.1. *$\xi_n(\theta_0)$ is asymptotically distributed according to the normal law $N_k(0, \Gamma(\theta_0))$.*

3. Results

Let $C_0(R^k)$ be a family of functions which are continuous on R^k and satisfy

$$\lim_{|h| \to \infty} f(h) = 0 .$$

Consider the uniform metric

$$\rho(f, g) = \sup\{|f(h) - g(h)|;\ h \in R^k\} .$$

Then, the metric space $(C_0(R^k), \rho)$ is complete and separable. Now, we define an extension of the likelihood ratio $Z_n(h)$ in (1.2) such that

$$
(3.1)\quad \bar{Z}_n(h)=\begin{cases} Z_n(h)\,, & \text{if } \theta_0+\dfrac{h}{\sqrt{n}}\in\Theta\,, \\ \prod_{i=1}^{n}\left\{g\left(X_{i-1}, X_i\,;\,\theta_0+\dfrac{h}{\sqrt{n}}, 0\right)\Big/f(X_{i-1}, X_i\,;\,\theta_0)\right\}\,, & \text{if } \theta_0+\dfrac{h}{\sqrt{n}}\in\bar{\Theta}\,, \\ 0\,, & \text{if } \delta\left(\theta_0+\dfrac{h}{\sqrt{n}}, \bar{\Theta}\right)\geqq\dfrac{h}{\sqrt{n}}\,, \\ \text{continuous and } 0\leqq|\bar{Z}_n(h)|\leqq|Z_n(h)|, & \text{otherwise}\,. \end{cases}
$$

The results of this paper are the following theorems. Suppose that Assumptions A, B and C hold.

THEOREM 3.1. *The finite-dimensional distributions of* $h\rightleftarrows\bar{Z}_n(h)$ *converge to the corresponding those of* $h\rightleftarrows Z(h)$ *where*

$$
(3.2)\qquad Z(h)=\exp\left\{h^t\Gamma(\theta_0)^{1/2}\xi-\frac{1}{2}h^t\Gamma(\theta_0)h\right\}.
$$

and ξ *is the k-dimensional standard normal random vector* $N_k(0, I)$.

THEOREM 3.2. *Under Assumptions* A, B *and* C, *there exist positive constants* c_{01} *and* $c_{02}>0$ *such that*

$$
(3.3)\qquad \mathrm{P}\{\sup_{l\leq|h|\leq l+1} Z_n(h)>e^{-c_{01}l^2}\}\leqq c_{02}/l^2 \qquad (\text{integer } l\geqq 1)\,,
$$

$$
(3.4)\qquad \mathrm{P}\{\sup_{|h|\geq M} Z_n(h)>e^{-c_{01}M^2}\}\leqq c_{02}/M \qquad (\text{integer } M\geqq 1)\,.
$$

THEOREM 3.3. *The sample function of* $\bar{Z}_n(h)$ *belongs to* $C_0(R^k)$ *with probability one.*

THEOREM 3.4. *For any* $\varepsilon>0$

$$
\lim_{d\to 0}\overline{\lim_{n\to\infty}}\,\mathrm{P}\{\sup_{|h_1-h_2|<d}|\bar{Z}_n(h_1)-\bar{Z}_n(h_2)|>\varepsilon\}=0\,.
$$

Since $\bar{Z}_n(h)\in C_0(R^k)$ with probability one (by Theorem 3.3), we can regard $\{\bar{Z}_n(h)\}$ to be continuous functions on a compact metric space $(\bar{R}^k, \delta)$. Thus, on account of Ascoli-Alzera's Theorem tightness of $\{\bar{Z}_n(h)\}$ is equivalent to the assertion of Theorem 3.4 together with $\bar{Z}_n(0)=1$. (For example, see Billingsley [3] and Straf [11].) That is:

THEOREM 3.5. *The family of random fields,* $\{\bar{Z}_n(h)\}_{n=1,2,\dots}$ *is tight.*

Theorems 3.1 and 3.5 immediately conclude the following two theorems. (See Billingsley [3], Section 5 for the proofs.)

THEOREM 3.6. *The likelihood ratio random fields* $h \rightleftarrows \bar{Z}_n(h)$ *weakly converge to the random field* $h \rightleftarrows Z(h)$ *defined in* (3.2), *as* $n \to \infty$.

THEOREM 3.7. *For any measurable functionals* $\{\psi_n\}$ *on* $C_0(R^k)$ *which continuously converge to* ψ, *random variable* $\psi_n(\bar{Z}_n)$ *converges in distribution to* $\psi(Z)$, *as* $n \to \infty$.

4. The proofs of theorems

The following two lemmas are provided for the sake of the proof of Theorem 3.1.

LEMMA 4.1. *Suppose that Assumptions* A *and* B *hold. Consider* $\xi_n(\theta)$ *defined in* (2.29). *Then, for any positive* $M>0$,

$$\sup_{|h|\leq M} |\xi_n(\theta_0+h/\sqrt{n})-\xi_n(\theta_0)+\Gamma(\theta_0)h| \to 0 \,,$$

in probability, as $n \to \infty$.

PROOF. For any ε with $0<\varepsilon\leqq 1$, we can choose a positive integer n_0 and a positive number d such that

$$(4.1)\quad \begin{aligned} & M/\sqrt{n_0}<d_0 \,, \\ & |\sqrt{n}\,\lambda(\theta_0+h/\sqrt{n})+\Gamma(\theta_0)h|<\varepsilon/4 \qquad \text{for } |h|\leqq M \text{ and } n\geqq n_0 \\ & 0<H_1 d^2<\varepsilon/4 \,, \\ & |\Gamma(\theta_0)|\,d<\varepsilon/4 \,, \end{aligned}$$

because of Assumptions (B5) and (B6). The region $\{|h|\leqq M\}$ is covered by finite open sets

$$W(h_s)=\{h;\ |h-h_s|<d\}\,, \qquad s=1,\cdots,m\,.$$

Then, we obtain the following inequalities.

$$(4.2)\quad \begin{aligned} & \sup_{|h|\leq M} |\xi_n(\theta_0+h/\sqrt{n})-\xi_n(\theta_0)+\Gamma(\theta_0)h| \\ & \leqq \max_{1\leq s\leq m}\Bigg[\frac{d}{\sqrt{n}}\sum_{i=1}^{n}\{u(X_{i-1}, X_i;\theta_0+h_s/\sqrt{n}, d/\sqrt{n}) \\ & \quad -\mathrm{E}\,u(X_{i-1}, X_i;\theta_0+h_s/\sqrt{n}, d/\sqrt{n})\} \\ & \quad +\Bigg|\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\{\eta(X_{i-1}, X_i;\theta_0+h_s/\sqrt{n})-\eta(X_{i-1}, X_i;\theta_0) \end{aligned}$$

$$-\lambda(\theta_0+h_s/\sqrt{n})\}\Big|\Big]+3\varepsilon/4\,, \qquad \text{for } n\geqq n_0\,.$$

Therefore by Chebyshev's inequality and Lemma 2.2 together with Assumptions (A6), (B6), (B7) and (2.8), we have that

$$\text{(4.3)}\qquad P\{\sup_{|h|\leqq M}|\xi_n(\theta_0+h/\sqrt{n})-\xi_n(\theta_0)+\Gamma(\theta_0)h|>\varepsilon\}$$
$$\leqq n^{-1/2q}m\left\{1+4\sum_{i=1}^{\infty}\phi(i-1)^{1/p}\right\}(8^2/\varepsilon^2)(H_{2q})^{1/q}\{d^{2+1/q}+M^{1/q}\}\to 0\,,$$
$$\text{as } n\to\infty\,.$$

This completes the proof.

Set

$$\text{(4.4)}\qquad L_n(h)=\sum_{i=1}^{n}\log\{f(X_{i-1},X_i;\theta_0+h/\sqrt{n})/f(X_{i-1},X_i;\theta_0)\}\,.$$

From (1.2), $L_n(h)=\log\{Z_n(h)\}$. The following lemma is proved in the same way as the proof of Lemma 2.2 in Inagaki and Ogata [7].

LEMMA 4.2. *Under the same assumptions in Lemma* 4.1, *for any* $M>0$,

$$\sup_{|h|\leqq M}\left|L_n(h)-h^t\xi_n(\theta_0)+\frac{1}{2}h^t\Gamma(\theta_0)h\right|\to 0$$

in probability, as $n\to\infty$.

THE PROOF OF THEOREM 3.1. By Theorem 2.1,

$$\text{(4.5)}\qquad \xi_n(\theta_0)\to\Gamma(\theta_0)^{1/2}\xi$$

in law, where ξ is distributed to the k-dimensional standard normal distribution $N_k(0, I)$. Let

$$L_n^*(h)=L_n(h)+\frac{1}{2}h^t\Gamma(\theta_0)h=\log\left\{Z_n(h)\exp\frac{1}{2}h^t\Gamma(\theta_0)h\right\}\,.$$

Then, by Lemma 4.2, we have that for any $h_1,\cdots,h_m$ with $\theta_0+h_i/\sqrt{n}\in\Theta$, $i=1,\cdots,m$ and any real numbers $a_1,\cdots,a_m$,

$$\text{(4.6)}\qquad a_1L_n^*(h_1)+\cdots+a_mL_n^*(h_m)-(a_1h_1+\cdots+a_mh_m)^t\xi_n(\theta_0)\to 0$$
$$\text{in probability}\,.$$

Thus, from (4.5) and (4.6) we complete the proof of Theorem 3.1.

The following four lemmas are provided for the sake of the proof of Theorem 3.2.

LEMMA 4.3. *Under Assumptions* A *and* B, *there exist positive numbers,* d_1, $0<d_1\leqq d_0$, c_1 *and* $c_2>0$ *such that for any* h, $|h/\sqrt{n}|<d_1$

$$\mathrm{P}\{Z_n(h)>\exp(-c_1|h|^2)\}\leqq c_2/|h|^2 .$$

PROOF. It follows from Assumptions B and Remark (a) that

$$nK(\theta_0+\delta h)=\delta^2 h^t\left[\int_0^1 du\int_0^u -\Lambda(\theta_0+v\delta h)dv\right]h$$

and hence, that there exists a positive number d, $0<d\leqq d_0$, such that for h, $|h/\sqrt{n}|<d$,

$$\left|nK(\theta_0+h/\sqrt{n})-\frac{1}{2}h^t\Gamma(\theta_0)h\right|<\frac{1}{8}\gamma_0|h|^2 \tag{4.7}$$

where γ_0 is a positive number satisfying

$$\gamma_0|h|^2\leqq h^t\Gamma(\theta_0)h . \tag{4.8}$$

Similarly, it follows from Assumptions B that

$$\mathrm{E}\left|\frac{1}{\sqrt{n}}h^t\int_0^1\eta(X_0, X_1; \theta_0+uh/\sqrt{n})du\right|^{2q}$$
$$\leqq(|h|/\sqrt{n})^{2q}\{\mathrm{E}\,u(X_0, X_1; \theta_0, |h|/\sqrt{n})^{2q}+\mathrm{E}|\eta(X_0, X_1; \theta_0)|^{2q}\} ,$$

and hence, that for any h, $|h|/\sqrt{n}<d_0$,

$$\begin{aligned}&\mathrm{E}|\log\{f(X_0, X_1; \theta_0+h/\sqrt{n})/f(X_0, X_1; \theta_0)\}|^{2q}\\&\leqq(|h|/\sqrt{n})^{2q}\{H_{2q}d_0+H_0\} .\end{aligned} \tag{4.9}$$

Recall $L_n(h)=\log Z_n(h)$ and $\mathrm{E}\,L_n(h)=-nK(\theta_0+h/\sqrt{n})$. From (4.7) and (4.8), we have that for h, $|h/\sqrt{n}|<d$,

$$nK(\theta_0+h/\sqrt{n})>\frac{1}{2}\gamma_0|h|^2-\frac{1}{8}\gamma_0|h|^2=\frac{1}{4}\gamma_0|h|^2 . \tag{4.10}$$

Thus, we have from (4.10) and by Chebyshev's Inequality that

$$\begin{aligned}&\mathrm{P}\left\{Z_n(h)>\exp\left(-\frac{1}{8}\gamma_0|h|^2\right)\right\}\\&\leqq\mathrm{P}\left\{L_n(h)-\mathrm{E}\,L_n(h)>\frac{1}{8}\gamma_0|h|^2\right\}\\&\leqq\frac{1}{\{(1/8)\gamma_0|h|^2\}^2}\mathrm{E}|L_n(h)-\mathrm{E}\,L_n(h)|^2 .\end{aligned} \tag{4.11}$$

By Lemma 2.2 and (4.9), we have that

$$\mathrm{E}|L_n(h)-\mathrm{E}\,L_n(h)|^2 \tag{4.12}$$

$$\leqq n[\mathrm{E}|\log\{f(X_0, X_1; \theta_0+h/\sqrt{n})/f(X_0, X_1; \theta_0)\}|^{2q}]^{1/q}$$
$$\times\left[1+4\sum_{i=1}^{n-1}\phi(i-1)^{1/p}\right]$$
$$\leqq |h|^2\{H_{2q}d_0+H_0\}^{1/q}\left[1+4\sum_{i=1}^{n-1}\phi(i-1)^{1/p}\right].$$

Thus, from (4.11) and (4.12) we have that

$$\mathrm{P}\left\{Z_n(h)>\exp\left(-\frac{1}{8}\gamma_0|h|^2\right)\right\}$$
$$\leqq \frac{1}{|h|^2}\frac{64}{\gamma_0^2}\{H_{2q}d_0+H_0\}^{1/q}\left[1+4\sum_{i=1}^{n-1}\phi(i-1)^{1/p}\right].$$

Choose $c_1=\frac{1}{8}\gamma_0$ and $c_2=\frac{64}{\gamma_0^2}\{H_{2q}d_0+H_0\}^{1/q}\left[1+4\sum_{i=1}^{\infty}\phi(i-1)^{1/p}\right]$. Then the proof of this lemma is complete.

LEMMA 4.4. *Suppose the same assumptions as in Lemma* 4.3. *For* c_1 *and* d_1 *chosen in Lemma* 4.3, *there exists a positive constant* $c_3>0$ *such that for any integer* l, $1\leqq l$, $l+1\leqq d_1\sqrt{n}$,

$$\mathrm{P}\left\{\sup_{l\leqq|h|\leqq l+1} Z_n(h)>\exp\left(-\frac{c_1}{2}l^2\right)\right\}\leqq c_3/l^2.$$

Proof of the lemma is performed in parallel to that of Lemma 3.2 in [7] by applying Lemma 2.2. See [8] for the detail.

LEMMA 4.5. *Suppose Assumptions* A *and* C *hold. Then, for any* d *and* $M>0$, *there exist positive numbers* c_4 *and* $c_5>0$ *such that for any integer* l *with* $d\sqrt{n}\leqq l\leqq M\sqrt{n}$

$$\mathrm{P}\{\sup_{l\leqq|h|\leqq M\sqrt{n}} Z_n(h)>e^{-c_4l^2}\}<c_5/l^2.$$

PROOF. Set $\Theta_1=\{\theta\in\bar{\Theta};\ d\leqq|\theta-\theta_0|\leqq M\}$. Then, Θ_1 is compact because of the compactness of $\bar{\Theta}$. It follows from (2.18) of Remark (f) that for $\theta\in\Theta_1$ there is $d(\theta)>0$ satisfying

$$\mathrm{E}\log\{g(X_0, X_1; \theta, d(\theta))/f(X_0, X_1; \theta_0)\}<-\frac{1}{2}K(\theta)<0. \tag{4.13}$$

According to the compactness of Θ_1, there are finite numbers of points $\theta_1,\cdots,\theta_m$ such that $\Theta_1\subset\bigcup_{i=1}^{m}U_{d(\theta_s)}(\theta_s)$. This and (4.13) imply that

$$\bar{K}=\min\{\bar{K}(\theta_s);\ s=1,\cdots,m\}\ \text{(say)}\ >0$$

and hence that

$$-\mathrm{E}\log\{g(X_0, X_1; \theta_s, d(\theta_s))/f(X_0, X_1; \theta_0)\} - \frac{1}{4}\bar{K} > \frac{1}{4}\bar{K} > 0\,,$$
$$s=1,\cdots,m\,,$$

Choose $c_4=\bar{K}/(4M^2)$. Since for $d\sqrt{n}\leqq l\leqq M\sqrt{n}$

$$c_4l^2\leqq c_4M^2n=\frac{1}{4}\bar{K}n\,, \tag{4.14}$$

We have that for $d\sqrt{n}\leqq l\leqq M\sqrt{n}$

$$\begin{aligned}
&\mathrm{P}\{\sup_{l\leqq|h|\leqq M\sqrt{n}} Z_n(h)\geqq e^{-c_4l^2}\}\\
&\leqq \mathrm{P}\left[\sup_{s=1,\cdots,m}\sum_{i=1}^{n}\log\{g(X_{i-1}, X_i; \theta_s, d(\theta_s))/f(X_{i-1}, X_i; \theta_0)\}\geqq -\frac{1}{4}\bar{K}n\right]\\
&\leqq \sum_{s=1}^{m}\frac{1}{((1/4)\bar{K}n)^2}n(\mathrm{E}|\log\{g(X_0, X_1; \theta_s, d(\theta_s))/f(X_0, X_1; \theta_0)\}|^{2q})^{1/q}\\
&\quad\times\left\{1+4\sum_{i=1}^{n}\phi(i-1)^{1/p}\right\}
\end{aligned}$$

(by Chebyshev's Inequality and Lemma 2.2)

$$\leqq\frac{1}{l^2}\frac{4m}{c_4\bar{K}}(\bar{K}_1)^{1/q}\left\{1+4\sum_{i=1}^{\infty}\phi(i-1)^{1/p}\right\}\qquad\text{(from (4.16))}$$

where

$$\bar{K}_1=\max_{s=1,\cdots,m}\mathrm{E}|\log\{g(X_0, X_1; \theta_s, d(\theta_s))f(X_0, X_1; \theta_0)\}|^{2q}\qquad\text{(say)}$$

and Assumption (C3) means $0<\bar{K}_1<\infty$. This completes the lemma.

LEMMA 4.6. *There exist positive constants M_0, c_6 and $c_7>0$ such that for any integer $l\geqq M_0\sqrt{n}$*

$$\mathrm{P}\{\sup_{|h|\geqq l} Z_n(h)>e^{-c_6l^2}\}\leqq c_7/l^2\,.$$

PROOF. Since from (2.9) and (2.10)

$$\begin{aligned}
g(x_0, x_1; \theta_\infty, d)&=\sup\{f(x_0, x_1; \theta);\ |\theta-\theta_0|>(1-d)/d\}\\
&\geqq\sup\{f(x_0, x_1; \theta);\ |\theta-\theta_0|>1/d\}\,,\qquad\text{for } 0<d<1\,,
\end{aligned}$$

we see that

$$\sup_{|h|\geqq l} Z_n(h)\leqq\prod_{i=1}^{n}\{g(X_{i-1}, X_i; \theta_\infty, \sqrt{n}/l)/f(X_{i-1}, X_i; \theta_0)\}\,. \tag{4.15}$$

From Assumption (C4) it follows that there exist positive numbers M_0, $\bar{K}_2$ and $c_6>0$ such that for any integer $l\geqq M_0\sqrt{n}$,

$$(4.16)\quad \begin{aligned} &\mathrm{E}\log\{f(X_0, X_1; \theta_0)/g(X_0, X_1; \theta_\infty, \sqrt{n}/l)\} > 2c_6 l^2/n\,, \\ &\mathrm{E}|\log\{g(X_0, X_1; \theta_\infty, \sqrt{n}/l)/f(X_0, X_1; \theta_0)\} \\ &\quad -\mathrm{E}\log\{g(X_0, X_1; \theta_\infty, \sqrt{n}/l)/f(X_0, X_1; \theta_0)\}|^{2q} < \bar{K}_2(l/\sqrt{n})^2\,. \end{aligned}$$

Thus, using Chebyshev's Inequality and Lemma 2.2 we have from (4.15) and (4.16) that for any integer $l \geqq M_0\sqrt{n}$,

$$\begin{aligned} &\mathrm{P}\{\sup_{|h|\geqq l} Z_n(h) \geqq e^{-c_6 l^2}\} \\ &\leqq \mathrm{P}\Big[\sum_{i=1}^{n}[\log\{g(X_{i-1}, X_i; \theta_\infty, \sqrt{n}/l)/f(X_{i-1}, X_i; \theta_0)\} \\ &\qquad -\mathrm{E}\log\{g(X_{i-1}, X_i; \theta_\infty, \sqrt{n}/l)/f(X_{i-1}, X_i; \theta_0)\}] \geqq c_6 l^2\Big] \\ &\leqq \frac{1}{l^2}\frac{1}{c_6^2}(\bar{K}_2)^{1/q}\Big\{1+4\sum_{i=1}^{\infty}\phi(i-1)^{1/p}\Big\}\,. \end{aligned}$$

This completes the lemma.

THE PROOF OF THEOREM 3.2. The inequality (3.3) is an immediate result of Lemmas 4.4, 4.5 and 4.6. Since

$$\begin{aligned} \mathrm{P}\{\sup_{|h|\geqq M} Z_n(h) > e^{-c_{01}M^2}\} &\leqq \mathrm{P}\Big\{\sup_{|h|>M} Z_n(h) > \sum_{l=M}^{\infty} e^{-c_{01}l^2}\Big\} \\ &\leqq \sum_{l=M}^{\infty}\mathrm{P}\{\sup_{l\leqq|h|\leqq l+1} Z_n(h) > e^{-c_{01}l^2}\}\,, \end{aligned}$$

it follows from (3.3) that

$$\mathrm{P}\{\sup_{|h|>M} Z_n(h) > e^{-c_{01}M^2}\} \leqq c_{02}\sum_{l=M}^{\infty}\frac{1}{l^2} \leqq 2c_{02}\frac{1}{M}\,, \qquad \text{for } M \geqq 1\,.$$

This leads Theorem 3.2.

THE PROOF OF THEOREM 3.3. (3.4) in Theorem 3.2 and the definition (3.1) of $\bar{Z}_n(h)$ imply Theorem 3.3.

THE PROOF OF THEOREM 3.4. By the definition (3.1) of $\bar{Z}_n$ it is sufficient to show that the assertion of Theorem 3.4 holds with respect to Z_n (in the place of $\bar{Z}_n$).

Choose c_{01} and $c_{02}>0$ such as in Theorem 3.2 and $M_1>0$ such that

$$(4.17)\qquad e^{-c_{01}M_1^2} < \varepsilon' \quad\text{and}\quad c_{02}/M_1 < \varepsilon\,.$$

Then, it follows from (3.4) in Theorem 3.2 and (4.19) that

$$(4.18)\quad \begin{aligned} &\mathrm{P}[\sup\{|Z_n(h_1)-Z_n(h_2)|;\ |h_1-h_2|<d \text{ and } |h_1|, |h_2| \geqq M_1\} > \varepsilon'] \\ &\leqq \mathrm{P}[\sup\{|Z_n(h)|;\ |h| \geqq M_1\} > \varepsilon'] < \varepsilon\,. \end{aligned}$$

Now, let $M_2 \geqq M_1+1$ (and $d \leqq 1$). Since $e^x - e^y = \int_y^x e^t dt$ and $Z_n(h) = e^{L_n(h)}$, we have that

$$
\begin{aligned}
(4.19)\quad & \sup\{|Z_n(h_1)-Z_n(h_2)|;\ |h_1-h_2|<d \text{ and } |h_1|, |h_2| \leqq M_2\} \\
& \leqq \sup\{Z_n(h);\ |h| \leqq M_2\} \\
& \quad \times \sup\{|L_n(h_1)-L_n(h_2)|;\ |h_1-h_2|<d \text{ and } |h_1|, |h_2| \leqq M_2\}\,.
\end{aligned}
$$

Further we have that

$$
\begin{aligned}
(4.20)\quad & \sup\{Z_n(h);\ |h| \leqq M_2\} \\
& \leqq \exp\left\{\sup_{|h| \leqq M_2}\left|L_n(h)-h^t\xi_n(\theta_0)+\frac{1}{2}h^t\Gamma(\theta_0)h\right| \right. \\
& \qquad \left. + M_2|\xi_n(\theta_0)| + \frac{1}{2}M_2^2|\Gamma(\theta_0)|\right\}\,,
\end{aligned}
$$

and

$$
\begin{aligned}
(4.21)\quad & \sup\{|L_n(h_1)-L_n(h_2)|;\ |h_1-h_2|<d \text{ and } |h_1|, |h_2| \leqq M_2\} \\
& \leqq 2\sup_{|h| \leqq M_2}\left|L_n(h)-h^t\xi_n(\theta_0)+\frac{1}{2}h^t\Gamma(\theta_0)h\right| \\
& \quad + d|\xi_n(\theta_0)| + dM_2|\Gamma(\theta_0)|\,.
\end{aligned}
$$

It follows from (4.19), (4.20) and (4.21) that

$$
\begin{aligned}
(4.22)\quad & \sup\{|Z_n(h_1)-Z_n(h_2)|;\ |h_1-h_2|<d \text{ and } |h_1|, |h_2| \leqq M_2\} \\
& \leqq \left[\exp\left\{\sup_{|h| \leqq M_2}\left|L_n(h)-h^t\Gamma_n(\theta_0)+\frac{1}{2}h^t\Gamma(\theta_0)h\right| \right.\right. \\
& \qquad \left.\left. + M_2|\xi_n(\theta_0)| + \frac{1}{2}M_2^2|\Gamma(\theta_0)|\right\}\right] \\
& \quad \times [2\sup_{|h| \leqq M_2}|L_n(h)-h^t\xi_n(\theta_0)+h^t\Gamma(\theta_0)h| \\
& \qquad + d|\xi_n(\theta_0)| + dM_2|\Gamma(\theta_0)|]\,.
\end{aligned}
$$

Therefore it follows that

$$
\begin{aligned}
(4.23)\quad & \mathrm{P}\,[\sup\{|Z_n(h_1)-Z_n(h_2)|>\varepsilon;\ |h_1-h_2|<d \text{ and } |h_1|, |h_2| \leqq M_2\} > \varepsilon'] \\
& \leqq \mathrm{P}\left\{\sup_{|h| \leqq M_2}\left|L_n(h)-h^t\xi_n(\theta_0)+\frac{1}{2}h^t\Gamma(\theta_0)h\right| > \varepsilon'\right\} \\
& \quad + \mathrm{P}\,\{|\xi_n(\theta_0)|>a\} \\
& \quad + \mathrm{P}\left[2\sup_{|h| \leqq M_2}\left|L_n(h)-h^t\xi_n(\theta_0)+\frac{1}{2}h^t\Gamma(\theta_0)h\right| \right. \\
& \left. > -d(a+M_2|\Gamma(\theta_0)|)+\varepsilon'\exp\left\{-\left(\varepsilon'+M_2a+\frac{1}{2}M_2^2|\Gamma(\theta_0)|\right)\right\}\right]\,.
\end{aligned}
$$

By Chebyshev's Inequality and Lemma 2.2 together with Assumption

(B7), (2.15) and (2.29), we have that

$$\text{(4.24)} \qquad \begin{aligned} \mathrm{P}\,\{|\xi_n(\theta_0)|>a\} &= \mathrm{P}\left\{\left|\sum_{i=1}^{n}\eta(X_{i-1}, X_i; \theta_0)\right|>\sqrt{n}\,a\right\} \\ &\leqq \frac{1}{a^2}kH_0^{1/q}\left\{1+4\sum_{i=1}^{\infty}\phi(i-1)\right\}<\varepsilon \end{aligned}$$

choosing $a>0$ so large for $\varepsilon>0$. By Lemma 4.2 we have that

$$\text{(4.25)} \qquad \mathrm{P}\left\{\sup_{|h|\leqq M_2}\left|L_n(h)-h^t\xi_n(\theta_0)-\frac{1}{2}h^t\Gamma(\theta_0)h\right|>\varepsilon''\right\}\leqq\varepsilon$$

choosing n so large for ε'' and $\varepsilon>0$. Now for sufficiently small $d>0$, let

$$\text{(4.26)} \qquad \begin{aligned} \varepsilon''=\min\Big[\varepsilon', &-d(a+M_2|\Gamma(\theta_0)|) \\ &+\varepsilon'\exp\left\{-\left(\varepsilon'+M_2a+\frac{1}{2}M_2^2|\Gamma(\theta_0)|\right)\right\}\Big]\,. \end{aligned}$$

Then we can take $\varepsilon''>0$. After all, it follows from (4.23)-(4.26) that for any ε and $\varepsilon'>0$ there exist n_0 and $d_0'>0$ such that for any $n\geqq n_0$ and d, $0<d<d_0'$,

$$\text{(4.27)} \quad \mathrm{P}\,[\sup\{|Z_n(h_1)-Z_n(h_2)|;\ |h_1-h_2|<d \text{ and } |h_1|, |h_2|\leqq M_2\}>\varepsilon']\leqq 3\varepsilon\,.$$

Thus from (4.18) and (4.27) we have the conclusion of Theorem 3.4.

5. Statistical Markovian models

Let a sequence of random vectors $\{X_n\}$ be generated by the relation

$$\text{(5.1)} \qquad X_n=AX_{n-1}+Y_n\,, \qquad n=1, 2, \cdots,$$

where the elements of the $p\times p$-matrix $A=(\alpha_{ij})$ are dominated by q-dimensional parameter $\theta=(\theta_1, \cdots, \theta_q)$, that is $\alpha_{ij}=\alpha_{ij}(\theta_1, \cdots, \theta_q)$, such that all the eigenvalues of the matrix are within a unit circle. Random vectors $\{Y_n\}$ are identically independent distributed according to a probability density function $f(y_1, \cdots, y_p)$ with respect to some σ-finite measure $\mu(dy_1, \cdots, dy_p)$, and further Y_n is independent of X_m for all $m\leqq n-1$. Then the density function of the transition probability with respect to the σ-finite measure μ is given for vectors ξ, η,

$$\text{(5.2)} \qquad f(\xi, \eta; \theta)=f(\eta-A\xi)\,.$$

Therefore the Fisher's information matrix (2.5) is given with the elements for $i, j=1, \cdots, q$.

(5.3) $$\gamma_{ij}(\theta)=\text{trace}\left\{\Lambda\frac{\partial A}{\partial\theta_i}\Sigma\left(\frac{\partial A}{\partial\theta_j}\right)^t\right\},$$

where Σ is the solution of the equation

$$\Sigma=\Sigma_0+A^t\Sigma A$$

for a given covariance matrix Σ_0 of the random vector Y_n, and matrix

$$\Lambda=(\lambda_{ij})=\left(\mathrm{E}\left\{\frac{\partial}{\partial y_i}\log f(Y_n)\frac{\partial}{\partial y_j}\log f(Y_n)\right\}\right)_{i,j=1,\cdots,p}.$$

Now we have the following result about the mixing coefficient (2.2).

THEOREM 5.1. *Suppose the density function of the random vector Y_n satisfies*

(5.4) $$\delta_1=\int_{R^p}|y|f(y)dy<\infty\ ,\qquad c_1=\int_{R^p}\left|\frac{\partial}{\partial y}f(y)\right|dy<\infty\ ,$$

where $y=(y_1,\cdots,y_p)$, and further the characteristic function $\phi(t)$, $t=(t_1,\cdots,t_p)$, of the random vector Y_n satisfies

(5.5) $$c_2=\int_{R^p}|t\phi(t)|dt<\infty\ .$$

Then $\phi(n)$ tends to zero with exponential order.

PROOF. The m-step transition density function of (5.1) is given by

(5.6) $$f^{(m)}(\xi,\eta;\theta)=f_m(\eta-A^m\xi;\theta)$$

where $f_m(\eta;\theta)$ is a density function of a random vector

(5.7) $$V_m=A^{m-1}Y_1+\cdots+AY_{m-1}+Y_m\ .$$

Let $\phi_m(t)$ be a characteristic function of the random vector V_m. Then we obtain

(5.8) $$\phi_m(t)=\mathrm{E}\left[\exp\{i(t,V_m)\}\right]=\prod_{k=0}^{n-1}\mathrm{E}\left[\exp\{i(t,A^kY_k)\}\right]=\prod_{k=0}^{n-1}\phi(tA^k)\ .$$

Since the absolute value of characteristic function is uniformly not larger than 1, we have

(5.9) $$1\geqq|\phi(t)|=|\phi_1(t)|\geqq|\phi_2(t)|\geqq\cdots\geqq|\phi_m(t)|\geqq\cdots\ .$$

Thus for any $n\geqq 1$

(5.10) $$\int|t\phi_m(t)|dt\leqq\int|t\phi(t)|dt<\infty\ .$$

By virtue of (5.10) we have the inverse formula

$$(5.11)\qquad f_m(\eta;\theta)=\frac{1}{2\pi}\int_{R^p}\psi_m(t)e^{-i(t,\eta)}dt\,,\qquad m=1,2,\cdots,$$

here $f_m(\eta;\theta)$ is the density function of the random vector of (5.7). Since all the eigenvalues of the matrix A are within a unit circle, there is a constant $0<\rho<1$ such that

$$(5.12)\qquad |A\xi|<\rho|\xi|\,,\qquad |\xi^t A|<\rho|\xi|$$

for all $\xi\in R^p$ and therefore the stationary initial density function is given

$$(5.13)\qquad f(\eta,\theta)=\lim_{m\to\infty}f_m(\eta;\theta)=\lim_{m\to\infty}f_m(\eta-A^m\xi;\theta)\qquad \text{(say)}\,.$$

Thus we have

$$(5.14)\qquad \begin{aligned}\phi(m)&=\iint|f_m(\eta-A^m\xi;\theta)-f(\eta;\theta)|f(\xi;\theta)d\xi d\eta\\ &\leqq\int|f_m(\eta;\theta)-f(\eta;\theta)|d\eta\\ &\quad+\iint|f(\eta-A^m\xi;\theta)-f(\eta;\theta)|f(\xi;\theta)d\xi d\eta\\ &=I_1+I_2\,.\end{aligned}$$

By the assumption (5.4) we obtain

$$(5.15)\qquad |\psi(t)-1|\leqq\int|(t,\xi)|f(\xi,\theta)d\xi\leqq\delta_1|t|\,.$$

Thus by (5.8), (5.10) and (5.12)

$$(5.16)\qquad \begin{aligned}&|f_m(\eta;\theta)-f_{m+1}(\eta;\theta)|\\ &\leqq\frac{1}{2\pi}\int|\psi_m(t)-\psi_{m+1}(t)|dt\leqq\frac{1}{2\pi}\int|\psi_m(t)||1-\psi(tA^m)|dt\\ &\leqq\frac{\delta_1}{2\pi}\int|tA^m||\psi_m(t)|dt<\frac{\delta_1}{2\pi}\rho^m\int|t||\psi_m(t)|dt=\frac{c_2\delta_1}{2\pi}\rho^m\,.\end{aligned}$$

Therefore for $N_m=\rho^{-m/2}$ we have

$$(5.17)\qquad \begin{aligned}I_1&\leqq\int\sum_{k=m}^{\infty}|f_k(\eta;\theta)-f_{k+1}(\eta;\theta)|d\eta\\ &\leqq\sum_{k=m}^{\infty}\left\{\int_{\{|\eta|\leqq N_m\}}|f_k(\eta;\theta)-f_{k+1}(\eta;\theta)|d\eta\right.\\ &\quad\left.+\int_{\{|\eta|>N_m\}}|f_k(\eta;\theta)-f_{k+1}(\eta;\theta)|d\eta\right\}\end{aligned}$$

$$\leqq \sum_{k=m}^{\infty} \left\{ \frac{c_2\delta_1}{2\pi} N_m \rho^m + \frac{2}{N_m} \int |\eta| f_k(\eta;\theta) d\eta \right\}$$

$$\leqq \sum_{k=m}^{\infty} \left\{ \frac{c_2\delta_1}{2\pi} \rho^{m/2} + \frac{2\delta_1}{1-\rho} \rho^{m/2} \right\} \leqq \frac{\delta_1}{1-\rho^{1/2}} \left(\frac{c_2}{2\pi} + \frac{2}{1-\rho} \right) \rho^{m/2} .$$

On the other hand from the assumption (5.5) we see that $f(y)$ has a bounded continuous derivative $(\partial/\partial y)f(y)$ such that

$$\left| \frac{\partial}{\partial y} f(y) \right| \to 0 \qquad \text{as } |y| \to \infty . \tag{5.18}$$

By (5.7) and (5.13) we have

$$f(\xi;\theta) = \int f(\xi - y) g(y;\theta) dy \tag{5.19}$$

for some probability density $g(y;\theta)$. Thus by virture of Fubini's theorem we obtain

$$\begin{aligned} \int \left| \frac{\partial}{\partial \xi} f(\xi;\theta) \right| d\xi &\leqq \iint \left| \frac{\partial}{\partial \xi} f(\xi - y) \right| g(y;\theta) d\xi dy \\ &= \int \left| \frac{\partial}{\partial \xi} f(\xi) \right| d\xi = c_1 < \infty . \end{aligned} \tag{5.20}$$

Therefore we have

$$\begin{aligned} I_2 &= \iint f(\xi;\theta) \left| \int_0^1 \frac{d}{d\beta} f(\eta - \beta A^m \xi;\theta) d\beta \right| d\xi d\eta \\ &= \int |A^m \xi| f(\xi;\theta) \int_0^1 \int \left| \frac{\partial}{\partial \eta} f(\eta - \beta A^m \xi;\theta) \right| d\eta d\beta d\xi \\ &\leqq \rho^m \int |\xi| f(\xi;\theta) d\xi \int \left| \frac{\partial}{\partial \eta} f(\eta;\theta) \right| d\eta \leqq \frac{c_1\delta_1}{1-\rho} \rho^m . \end{aligned} \tag{5.21}$$

This and (5.17) complete the proof.

We present several examples of Markovian statistical model where Theorem 5.1 can be applied.

Example 1 (Doeblin's condition). Let $\{X_n\}$ be ergodic and satisfy well-known Doeblin's condition (see Doob [5], p. 192, for example). Then

$$\sup_x |Q^{(n)}(x, B) - q(B)| \leqq C\rho^n , \qquad (\rho < 1) , \tag{5.22}$$

for any $B \in \mathfrak{B}$. This implies Assumption (A6) directly. It is also well-known that ergodic finite state Markov chain satisfies (5.22).

Example 2 (Gaussian simple autoregressive process). Let the chain be defined by the recurrence relation

(5.23) $$X_n = \theta X_{n-1} + \varepsilon_n\,, \qquad n=1, 2, \cdots,$$

where $\varepsilon_1, \varepsilon_2, \cdots$ are i.i.d. $N(0, \sigma^2)$ random variables and θ is a real number with

(5.24) $$|\theta| < 1\,.$$

Then the density of transition probability is given by

(5.25) $$f(x, y; \theta) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{1}{2\sigma^2}(y-\theta x)^2\right\}.$$

It is easily seen that the n-step transition probability density is given by

(5.26) $$f^{(n)}(x, y; \theta) = \frac{1}{\sqrt{2\pi}\sigma}\left(\frac{1-\theta^2}{1-\theta^{2n}}\right)^{1/2} \exp\left\{-\frac{1}{2\sigma^2}\left(\frac{1-\theta^2}{1-\theta^{2n}}\right)(y-\theta^n x)^2\right\}.$$

Thus, the density of the stationary initial distribution is as follows:

(5.27) $$f(x, \theta) = \frac{\sqrt{1-\theta^2}}{\sqrt{2\pi}\sigma} \exp\left(-\frac{1-\theta^2}{2\sigma^2}x^2\right).$$

Now, we have that

(5.28) $$\begin{aligned}\phi(n) &= \iint f(x, \theta)\,|f^{(n)}(x, y; \theta) - f(y; \theta)|\,dxdy \\ &\leqq 2\left\{\iint f(x, \theta)((f^{(n)}(x, y; \theta))^{1/2} - (f(y; \theta))^{1/2})^2 dxdy\right\}^{1/2} \\ &= 2\sqrt{2}\left\{1 - \iint f(x, \theta)(f^{(n)}(x, y; \theta))^{1/2}(f(y; \theta))^{1/2} dxdy\right\}^{1/2}.\end{aligned}$$

From (5.26) and (5.27) we see that

(5.29) $$\begin{aligned}&\iint f(x, \theta)(f^{(n)}(x, y; \theta))^{1/2}(f(y; \theta))^{1/2} dxdy \\ &\qquad = (1-\theta^{2n})^{3/4}/\{(1-\theta^{2n}/2)^2 - \theta^{2n}/4\}^{1/2}.\end{aligned}$$

Therefore we obtain from (5.8) and (5.9) that

(5.30) $$\phi(n) \leqq 2\sqrt{2}\,[1 - \{(1-\theta^{2n})^3/((1-\theta^{2n}/2)^2 - \theta^{2n}/4)^2\}^{1/4}]^{1/2}.$$

On account of the following inequality:

$$\{(1-a)^3/(1-5a/4+a^2/4)^2\}^{1/4} \geqq 1-3a\,, \qquad \text{for } 0 \leqq a \leqq 1/3,$$

we obtain from (5.28) and (5.29) that

$$\phi(n) \leqq 2\sqrt{6}\,\theta^n$$

for so large n that $\theta^{2n} \leqq 1/3$.

The last example does not, however, satisfy the Doeblin's condition in Example 1.

Example 3 (simple autoregressive process with stable distribution). Let the chain also be defined by the relation (5.23) where $\varepsilon_1, \varepsilon_2, \cdots,$ are i.i.d. random variables whose characteristic function is

$$\phi(t)=\exp[-c|t|^{\alpha}], \qquad c>0,\ 1<\alpha<2. \tag{5.31}$$

Though the random variables do not have a finite variance, the autoregressive process satisfies the assumption of Theorem 5.1.

Example 4 (simple autoregressive process with general distribution). Let the chain also be defined by the relation (5.23) where $\varepsilon_1, \varepsilon_2, \cdots,$ are i.i.d. random variables whose probability density function with respect to some σ-finite measure μ is given by $f(x)$. If $f(x)$ is three times differentiable in a real line, and

$$\begin{aligned} &\lim_{|x|\to\infty} f'(x)=\lim_{|x|\to\infty} f''(x)=0, \\ &c_3=\int_{-\infty}^{\infty}|f'''(x)|\mu(dx)<\infty, \end{aligned} \tag{5.32}$$

then this implies (5.5) for 1-dimensional case. In fact repeating the partial integration we obtain for the characteristic function

$$\phi(t)=-\int_{-\infty}^{\infty}\{f'''(x)e^{itx}/(it)^3\}\mu(dx).$$

Thus we have

$$|\phi(t)|\leqq\min\{1, c_3/t^3\}.$$

(5.4) and (5.32) implies the results of Theorem 5.1. Furthermore if ε_n has a finite variance $\sigma^2=\int_{-\infty}^{\infty}x^2f(x)\mu(dx)<\infty$, and then set

$$J=\int_{-\infty}^{\infty}\{f'(x)^2/f(x)\}\mu(dx), \tag{5.33}$$

thus Fisher's information is given by

$$\Gamma(\theta)=\sigma^2J/(1-\theta^2). \tag{5.34}$$

Example 5 (General autoregressive process). Let $\{X_n\}$ be an m-variate stationary autoregressive process generated by the relation

$$\sum_{k=0}^{p}A_kX_{n-k}=Y_n, \tag{5.35}$$

where the coefficients $A_k=(a_{ij}^{(k)})$ are $m\times m$-matrices with $A_0=I$, (identity).

The i.i.d. random vector Y_n has a density function $g(y_1)$ with respect to a σ-finite measure $\nu(dy_1)$. Consider the mp-dimensional vector $\zeta_n=(X_n, X_{n-1},\cdots, X_{n-p+1})'$, $\varepsilon_n=(Y_n, 0,\cdots, 0)'$ and $mp\times mp$-matrix

$$C=\begin{bmatrix} -A_1, & -A_2, \cdots, & -A_{p-1}, & -A_p \\ I & 0\ ,\cdots, & 0 & 0 \\ 0 & I\ ,\cdots, & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0\ ,\cdots, & I & 0 \end{bmatrix}. \tag{5.36}$$

Then the relation (5.35) is reduced to the relation (5.1), and the probability density function of ε_n is given $f(y_1,\cdots, y_p)=g(y_1)$ with respect to σ-finite measure $\mu(dy_1,\cdots, dy_p)=\nu(dy_1)$. Thus by virture of (5.3) and (5.36) Fisher's information matrix is given in the form of tensor product, that is,

$$\Gamma(\theta)=K\otimes\Sigma\ , \tag{5.37}$$

where $\Sigma=\mathrm{E}\,[\zeta_n\zeta_n']$ and $K=(\kappa_{ij})$ is $m\times m$-matrix such that

$$\kappa_{ij}=\mathrm{E}\left\{\frac{\partial}{\partial y_i}\log f(Y_n)\frac{\partial}{\partial y_j}\log f(Y_n)\right\}\ , \qquad i, j=1,\cdots, m\ . \tag{5.38}$$

THE INSTITUTE OF STATISTICAL MATHEMATICS
OSAKA UNIVERSITY

REFERENCES

[1] Bahadur, R. R. (1967). An optimal property of the likelihood ratio statistic, *Proc. 5th Berkeley Symp. Math. Statist. Prob.*, **1**, 13-36.

[2] Billingsley, P. (1961). The Lindeberg-Levy theorem for martingales, *Proc. Amer. Math. Soc.*, **12**, 788-792.

[3] Billingsley, P. (1968). *Convergence of Probability Measures*, John Wiley & Sons, Inc., New York.

[4] Davydov, Yu. A. (1973). Mixing conditions for Markov chains, *Theory Prob. Appl.*, **18**, 312-328.

[5] Doob, J. L. (1953). *Stochastic Processes*, John Wiley & Sons, Inc., New York.

[6] Ibragimov, I. A. and Khas'minskii, R. Z. (1972). Asymptotic behavior of statistical estimators in the smooth case, *Theory Prob. Appl.*, **17**, 443-460.

[7] Inagaki, N. and Ogata, Y. (1975). The weak convergence of likelihood ratio random fields and its applications, *Ann. Inst. Statist. Math.*, **27**, 391-419.

[8] Inagaki, N. and Ogata, Y. (1976). The weak convergence of the likelihood ratio random fields for Markov observations, *Research Memorandum*, No. 79, The Institute of Statistical Mathematics.

[9] LeCam, L. (1970). On the assumptions used to prove asymptotic normality of maximum likelihood estimates, *Ann. Math. Statist.*, **41**, 802-828.

[10] Ogata, Y. and Inagaki, N. (1976). The weak convergence of the likelihood ratio random fields for Markov observations II, (Asymptotic analysis for statistical Markovian models), *Research Memorandum*, No. 96, The Institute of Statistical Mathematics.

[11] Straf, M. L. (1970). Weak convergence of random processes with several parameters, *Proc. 6th Berkeley Symp. Math. Statist. Prob.*, **2**, 187-221.