数理解析研究所講究録
1055 巻 1998 年 79-95

On large deviation probability of sequential MLE for the exponential class

聖心女子大学　三田　晴義 (Haruyoshi Mita)

We investigate the asymptotic behavior of probability of large deviations for the sequential maximum likelihood estimator for processes of the exponential class with independent increments. It is shown that the probability of large deviations for the sequential maximum likelihood estimator decays exponentially fast as a stopping boundary diverges. Further we study the asymptotic efficiency of the sequential maximum likelihood estimator in the Bahadur sense. Many authors have studied the efficiency of sequential estimators in the decision theoretic sense. However, we try to study the asymptotic efficiency of the sequential maximum likelihood estimator in the sense of probability of large deviations.

1. The exponential class of processes with independent increments

Let $X(t), t \in T$, be a stochastic process defined on a probability space $(\Omega, \boldsymbol{F}, P_\theta)$, with values in $(R^1, \boldsymbol{B})$, where $T = [0,\infty)$ or $\{0,1,2,\cdots\cdots\}$ and $\boldsymbol{B}$ is the Borel $\sigma-$field in R^1. The probability P_θ depends on an unknown parameter $\theta \in \Theta$, where Θ is an open subset of R^1. Let $\boldsymbol{F}_t$, $t \in T$, be the $\sigma-$field generated by the process $X(s)$, $s \le t$, and the restriction of P_θ to the $\sigma-$field $\boldsymbol{F}_t$ is denoted by $P_{\theta,t}$.

Definition. The stochastic process $X(t)$ belongs to the *exponential class with independent increments*, if the following conditions are fulfilled:

(i) $X(t)$ is a stationary stochastic process with independent increments satisfying $P_\theta(X(0)=0)=1$ for all $\theta \in \Theta$ and continuous in probability.

(ii) The probability distributions at time t, that is, $P_{\theta,t}$, $\theta \in \Theta$ are dominated by the restriction of a probability measure μ on $\boldsymbol{F}_t$, which is denoted by μ_t, and the Radon-Nikodym derivatives may be represented in the form

$$p(x,t;\theta) := \frac{dP_{\theta,t}}{d\mu_t} = g(x,t)\exp(\theta x - f(\theta)t),$$

where g is a non-negative function defined on $R^1 \times T$ and f is a twice differentiable real valued function with $\ddot{f}(\theta) > 0$ for all $\theta \in \Theta$.

It is known that if a stochastic process $X(t)$ belongs to the exponential class with independent increments, then $X(t)$ is equivalent to a stochastic process $Y(t)$ having the property that almost all of its sample paths are right-continuous and have left-limits at each t, that is, have at most jump discontinuities of the first kind. Moreover, the process $Y(t)$ is unique in the sense that $\tilde{Y}(t)$ is any other such process, then $P(Y(t) = \tilde{Y}(t) \quad \text{for every } t) = 1$.

2. Stopped processes of the exponential class

Let τ be an arbitrary stopping time, that is, τ is a random variable defined on Ω with values in $T \cup \{\infty\}$ and has the property that $\{\omega \in \Omega. \tau(\omega) \le t\} \in \boldsymbol{F}_t$ for any $t \in T$. The $\sigma-$field of the $\tau-$past of the process $X(t)$ is denoted by

$$\boldsymbol{F}_\tau = \{F \in \boldsymbol{F}. F \cap \{\omega \in \Omega. \tau(\omega) \le t\} \in \boldsymbol{F}_t \text{ for any } t \in T\}.$$

We assume $P_\theta(\tau < \infty) = 1$ for any $\theta \in \Theta$. Let $P_{\theta,\tau}$ and μ_τ denote the restrictions of P_θ and μ on $\boldsymbol{F}_\tau$, respectively. It is known that $P_{\theta,\tau}$ is dominated by μ_τ and the corresponding likelihood function, which is denoted by $L_\tau(\theta)$, is represented as

$$L_\tau(\theta) := g(X(\tau), \tau)\exp(\theta X(\tau) - f(\theta)\tau). \tag{2.1}$$

See Basawa,I.V. and Prakasa Rao,B.L.S.(1980). This means that the likelihood function is independent of the sampling rule. Since $L_\tau(\theta)$ is the likelihood function of the exponential family, we obtain

$$E_\theta(X(\tau)) = \dot{f}(\theta)E_\theta(\tau). \tag{2.2}$$

3. Lower bounds for consistent estimators

To allow for asymptotic considerations, we introduce the stopping times indexed by the real parameter u. Let $\tau(u)$ be a stopping time indexed by $u \in \Gamma$, where Γ is either the set of non-negative real numbers or the set of non-negative integers. We consider the estimation of an unknown parameter $\theta \in \Theta$. Let $\varphi(\tau(u), X(\tau(u)))$ be an estimator of θ based on a sufficient statistic $(\tau(u), X(\tau(u)))$. For convenience we write $\varphi_{\tau(u)} = \varphi(\tau(u), X(\tau(u)))$. When a sequence of stopping times $\{\tau(u): u \in \Gamma\}$ is given, an estimator $\varphi_{\tau(u)}$ is said to be *consistent for θ with respect to the sequence of stopping times* $\{\tau(u): u \in \Gamma\}$ if for any $\theta \in \Theta$, $\varphi_{\tau(u)} \to \theta$ in probability as $u \to \infty$ under P_θ. Let $\boldsymbol{T}$ be a class of sequences of stopping times having the property that for any $\theta \in \Theta$, $\frac{\tau(u)}{u} \to c(\theta)$ in probability as $u \to \infty$ under P_θ and $\frac{E_\theta(\tau(u))}{u} \to c(\theta)$ as $u \to \infty$, where c is positive and continuous in Θ. Furthermore, let $\boldsymbol{C}$ be a class of estimators which are consistent for θ with respect to any stopping time sequence which belongs to $\boldsymbol{T}$. Let $K_u(\theta_1, \theta_2)$ be the Kullback-Leibler information distance from $P_{\theta_1,\tau(u)}$ to $P_{\theta_2,\tau(u)}$, that is, for any θ_1, $\theta_2 \in \Theta$,

$$K_u(\theta_1, \theta_2) := \int \log \frac{dP_{\theta_1,\tau(u)}}{dP_{\theta_2,\tau(u)}} dP_{\theta_1,\tau(u)}.$$

It follows that

$$\begin{aligned} K_u(\theta_1, \theta_2) &= (\theta_1 - \theta_2) E_{\theta_1}(X(\tau(u))) - (f(\theta_1) - f(\theta_2)) E_{\theta_1}(\tau(u)) \\ &= ((\theta_1 - \theta_2)\dot{f}(\theta_1) - (f(\theta_1) - f(\theta_2))) E_{\theta_1}(\tau(u)) \\ &= E_{\theta_1}(\tau(u)) K(\theta_1, \theta_2), \end{aligned}$$

where $K(\theta_1, \theta_2) = (\theta_1 - \theta_2)\dot{f}(\theta_1) - (f(\theta_1) - f(\theta_2))$.
We put $\tilde{K}(\theta_1, \theta_2) = \frac{c(\theta_1)}{c(\theta_2)} K(\theta_1, \theta_2)$.

Next Theorem gives us a lower bound for the probability of large deviation for any consistent estimator $\varphi_{\tau(u)}$ satisfying $\{\tau(u): u \in \Gamma\} \in \boldsymbol{T}$.

Theorem 1. *Suppose that* $\varphi_{\tau(u)}$ *is consistent for* θ *with respect to any sequence of stopping times satisfying* $\{\tau(u): u \in \Gamma\} \in \boldsymbol{T}$. *Then, for any sequence* $\{\tau(u): u \in \Gamma\} \in \boldsymbol{T}$ *it follows that for any* $\theta \in \Theta$ *and any* $\varepsilon > 0$ *satisfying* $\theta \pm \varepsilon \in \Theta$,

$$\liminf_{u \to \infty} \frac{1}{E_\theta(\tau(u))} \log P_{\theta,\tau(u)}(|\varphi_{\tau(u)} - \theta| > \varepsilon) \geq -B(\theta,\varepsilon),$$

where $B(\theta,\varepsilon) = \min\{\tilde{K}(\theta-\varepsilon,\theta), \tilde{K}(\theta+\varepsilon,\theta)\}$.

Proof. Fix $\{\tau(u): u \in \Gamma\} \in \boldsymbol{T}$ any. For any $\delta > 0$ and any $\varepsilon_1 > \varepsilon$ satisfying $\theta \pm \varepsilon_1 \in \Theta$, it follows that

$$\begin{aligned}
P_{\theta,\tau(u)}(|\varphi_{\tau(u)} - \theta| > \varepsilon) &= \int_{\{|\varphi_{\tau(u)}-\theta|>\varepsilon\}} dP_{\theta,\tau(u)} \\
&\geq \int_{\{|\varphi_{\tau(u)}-\theta|>\varepsilon\} \cap \left\{\frac{dP_{\theta+\varepsilon_1,\tau(u)}}{dP_{\theta,\tau(u)}} < e^\delta\right\}} \frac{dP_{\theta,\tau(u)}}{dP_{\theta+\varepsilon_1,\tau(u)}} dP_{\theta+\varepsilon_1,\tau(u)} \\
&\geq e^{-\delta} \int_{\{|\varphi_{\tau(u)}-\theta|>\varepsilon\} \cap \left\{\frac{dP_{\theta+\varepsilon_1,\tau(u)}}{dP_{\theta,\tau(u)}} < e^\delta\right\}} dP_{\theta+\varepsilon_1,\tau(u)} \\
&\geq e^{-\delta}\left(P_{\theta+\varepsilon_1,\tau(u)}\left(|\varphi_{\tau(u)} - \theta| > \varepsilon\right) - P_{\theta+\varepsilon_1,\tau(u)}\left(\frac{dP_{\theta+\varepsilon_1,\tau(u)}}{dP_{\theta,\tau(u)}} > e^\delta\right)\right). \qquad (3.1)
\end{aligned}$$

Since $\{\varphi_{\tau(u)}: u \in \Gamma\}$ is consistent for θ, we have

$$\lim_{u \to \infty} P_{\theta+\varepsilon_1,\tau(u)}\left(|\varphi_{\tau(u)} - \theta| > \varepsilon\right) = 1. \qquad (3.2)$$

Let $\delta = E_\theta(\tau(u))(K(\theta+\varepsilon_1,\theta) + \delta_1)\dfrac{c(\theta+\varepsilon)}{c(\theta)}$, where $\delta_1 > 0$ is arbitrary. Then we have

$$\begin{aligned}
&P_{\theta+\varepsilon_1,\tau(u)}\left(\frac{dP_{\theta+\varepsilon_1,\tau(u)}}{dP_{\theta,\tau(u)}} > e^\delta\right) \\
&= P_{\theta+\varepsilon_1,\tau(u)}\left(\varepsilon_1 X(\tau(u)) - (f(\theta+\varepsilon_1) - f(\theta))\tau(u) > E_\theta(\tau(u))(K(\theta+\varepsilon_1,\theta) + \delta_1)\right) \\
&= P_{\theta+\varepsilon_1,\tau(u)}\left(\varepsilon_1 \frac{X(\tau(u))}{u} - (f(\theta+\varepsilon_1) - f(\theta))\frac{\tau(u)}{u} > \frac{E_\theta(\tau(u))}{u} \cdot \frac{c(\theta+\varepsilon)}{c(\theta)}(K(\theta+\varepsilon_1,\theta) + \delta_1)\right) \\
&= P_{\theta+\varepsilon_1,\tau(u)}\left(Y(\tau(u)) > \frac{E_\theta(\tau(u))}{u} \cdot \frac{c(\theta+\varepsilon)}{c(\theta)}(K(\theta+\varepsilon_1,\theta) + \delta_1)\right), \qquad (3.3)
\end{aligned}$$

where $Y(\tau(u)) = \varepsilon_1 \dfrac{X(\tau(u))}{u} - (f(\theta+\varepsilon_1) - f(\theta))\dfrac{\tau(u)}{u}$.

Since $X(t)$ belongs to the exponential class with independent increments, $\dfrac{X(t)}{t} \to \dot{f}(\theta+\varepsilon_1)$ with probability one as $t \to \infty$ under $P_{\theta+\varepsilon_1,\tau(u)}$. Since $\dfrac{\tau(u)}{u} \to c(\theta+\varepsilon_1) > 0$ in probability as $u \to \infty$ under $P_{\theta+\varepsilon_1,\tau(u)}$, it follows that $\tau(u) \to \infty$ in probability as $u \to \infty$ under $P_{\theta+\varepsilon_1,\tau(u)}$. Therefore, $\dfrac{X(\tau(u))}{\tau(u)} \to \dot{f}(\theta+\varepsilon_1)$ in probability as $u \to \infty$ under $P_{\theta+\varepsilon_1,\tau(u)}$. Hence, we have

$$\frac{X(\tau(u))}{u} \to c(\theta+\varepsilon_1)\dot{f}(\theta+\varepsilon_1) \quad \text{in probability as } u \to \infty \quad \text{under } P_{\theta+\varepsilon_1,\tau(u)}.$$

Therefore,

$$Y(\tau(u)) \to c(\theta+\varepsilon_1)K(\theta+\varepsilon_1,\theta) \text{ in probability as } u\to\infty \text{ under } P_{\theta+\varepsilon_1,\tau(u)}.$$

Further, it follows that

$$\frac{E_\theta(\tau(u))}{u}\cdot\frac{c(\theta+\varepsilon_1)}{c(\theta)}(K(\theta+\varepsilon_1,\theta)+\delta_1)\to c(\theta+\varepsilon_1)(K(\theta+\varepsilon_1,\theta)+\delta_1) \text{ as } u\to\infty.$$

By (3.3), it follows that

$$P_{\theta+\varepsilon_1,\tau(u)}\left(\frac{dP_{\theta+\varepsilon_1,\tau(u)}}{dP_{\theta,\tau(u)}}>e^{\delta}\right)\to 0 \quad \text{as } u\to\infty.$$

From (3.1) and (3.2), we have

$$\liminf_{u\to\infty}\frac{1}{E_\theta(\tau(u))}\log P_{\theta,\tau(u)}\left(|\varphi_{\tau(u)}-\theta|>\varepsilon\right)\ge -(K(\theta+\varepsilon_1,\theta)+\delta_1)\frac{c(\theta+\varepsilon_1)}{c(\theta)}.$$

Since $\delta_1>0$ and $\varepsilon_1>\varepsilon$ are arbitrary and c is continuous, we have

$$\liminf_{u\to\infty}\frac{1}{E_\theta(\tau(u))}\log P_{\theta,\tau(u)}\left(|\varphi_{\tau(u)}-\theta|>\varepsilon\right)\ge -K(\theta+\varepsilon,\theta)\frac{c(\theta+\varepsilon)}{c(\theta)}$$

$$= -\tilde{K}(\theta+\varepsilon,\varepsilon). \tag{3.4}$$

Replacing $\theta+\varepsilon_1$ by $\theta-\varepsilon_1$ in the above discussion, we obtain

$$\liminf_{u\to\infty}\frac{1}{E_\theta(\tau(u))}\log P_{\theta,\tau(u)}\left(|\varphi_{\tau(u)}-\theta|>\varepsilon\right)\ge -\tilde{K}(\theta-\varepsilon,\theta). \tag{3.5}$$

According to (3.4)and (3.5), the proof is completed. □

4. Bahadur efficiency for the sequential MLE

We introduce the following stopping time:

$$\tau_{\alpha,\beta}(u) := \inf\{t\colon \alpha X(t)+\beta t\ge u\}, \tag{4.1}$$

where $\alpha\neq 0$, β, and u>0 are constants, and α and β are chosen such that $P_\theta(\tau_{\alpha,\beta}(u)<\infty)=1$ for any $\theta\in\Theta$. We abbreviate the indices α and β, that is, we write $\tau(u)$ for $\tau_{\alpha,\beta}(u)$. Let $D_{\tau(u)}$ be the overshoot for the stopping time given by (4.1), that is,

$$D_{\tau(u)} := \alpha X(\tau(u))+\beta\tau(u)-u. \tag{4.2}$$

We have

$$\alpha E_\theta(X(\tau(u)))+\beta E_\theta(\tau(u)) = E_\theta(D_{\tau(u)}+u). \tag{4.3}$$

From (2.2), we have

$$(\alpha\dot{f}(\theta)+\beta)E_{\theta}(\tau(u)) = E_{\theta}(D_{\tau(u)}+u). \tag{4.4}$$

We define $h(\theta) := f(\theta)+\alpha^{-1}\beta\theta$. Since $P_{\theta}(\tau(u)\geq 0) = P_{\theta}(D_{\tau(u)}\geq 0) = 1$, we have

$$\alpha\dot{h}(\theta) = \alpha\dot{f}(\theta)+\beta > 0. \tag{4.5}$$

Hence, h is invertible on Θ. Since $\alpha\neq 0$, $X(\tau(u)) = \alpha^{-1}(u+D_{\tau(u)}-\beta\tau(u))$. Therefore, the likelihood function is represented as

$$L_{\tau(u)}(\theta) = g(X(\tau(u)),\tau(u))\exp(\theta\alpha^{-1}(u+D_{\tau(u)})-h(\theta)\tau(u)). \tag{4.6}$$

We denote $\phi_{X,\theta}(s), s\in R^1$, as the moment generating function of a random variable X under P_{θ}. Here we need the following assumption :

Assumption (A). For any $\theta\in\Theta$, there exist a neighborhood N_{θ} of θ and a random variable $M_{\theta}(u)$ having the property that for any $u\in\Gamma$,

(i) for any $\vartheta\in N_{\theta}$, $P_{\vartheta}(D_{\tau(u)}\leq M_{\theta}(u)) = 1$,

(ii) the distribution of $M_{\theta}(u)$ under P_{ϑ} is independent of u and $\vartheta\in N_{\theta}$,

and

(iii) the moment generating function of $M_{\theta}(u)$ under P_{θ} exists in a neighborhood of origin.

Assumption (A) is fulfilled for many stochastic processes including Wiener process, Poisson process, Bernoulli process, etc.

Now let $\hat{\theta}_{\tau(u)}$ be the maximum likelihood estimator for the stopped likelihood function $L_{\tau(u)}$. From (1), we have $\dot{f}(\hat{\theta}_{\tau(u)}) = \dfrac{X(\tau(u))}{\tau(u)}$. According to Sørensen (1986), we can show that if Assumption (A) is fulfilled then the maximum likelihood estimators $\hat{\theta}_{\tau(u)}$ is consistent for θ with respect to the sequence of the stopping times given by (4.1).

Now we consider the moment generating function of $\dfrac{\tau(u)}{u}$ under P_{θ}. Since $h(\theta)-s/u\in h(\Theta)$ for sufficiently large $u>0$, we have

$$\begin{aligned}\phi_{\tau(u)/u,\theta}(s) &= E_{\theta}(\exp((\tau(u)/u)s))\\ &= E_{\mu_{\tau(u)}}\left[g(X(\tau(u)),\tau(u))\exp\left(\alpha^{-1}\theta(u+D_{\tau(u)})-(h(\theta)-s/u)\tau(u)\right)\right]\end{aligned}$$

$$= \exp(\alpha^{-1}u(\theta - h^{-1}(h(\theta) - s/u))) \cdot E_{\mu_{\tau(u)}}[\exp(\alpha^{-1}(\theta - h^{-1}(h(\theta) - s/u))D_{\tau(u)})$$
$$\cdot g(X(\tau(u)), \tau(u)) \exp(\alpha^{-1}h^{-1}(h(\theta) - s/u)(u + D_{\tau(u)}) - (h(\theta) - s/u)\tau(u))]$$
$$= \exp(\alpha^{-1}u(\theta - h^{-1}(h(\theta) - s/u)))\phi_{D_{\tau(u)}, h^{-1}(h(\theta)-s/u)}(\alpha^{-1}(\theta - h^{-1}(h(\theta) - s/u))). \quad (4.7)$$

From a result of Sørensen(1986, Lemma 3.7), it follows that for any $\delta > 0$,

$$\phi_{D_{\tau(u)}, \theta'}(su^{-\delta}) \to 1 \text{ as } u \to \infty$$

uniformly for $(\theta', s) \in N_\theta \times [-s_1, s_1]$, where the interval $[-s_1, s_1]$ is contained in the domain of the moment generating function of $D_{\tau(u)}$ under P_θ. Therefore, by (4.5) we have

$$\phi_{\tau(u)/u,\, \theta}(s) \to \exp(\alpha^{-1}s/\dot{h}(\theta)) \text{ as } u \to \infty.$$

By the continuity theorem for moment generating functions and (4.5), we have

$$\tau(u)/u \to 1/\alpha\dot{h}(\theta) > 0 \text{ in probability as } u \to \infty \text{ under } P_\theta.$$

Thus, the first assertion has been shown.

By (4.4), (4.5), and Assumption (A), it follows that

$$E_\theta(\tau(u))/u = E_\theta(D_{\tau(u)}/u + 1)/\alpha\dot{h}(\theta) \to 1/\alpha\dot{h}(\theta) > 0 \text{ as } u \to \infty.$$

Therefore we obtain the following result:

Lemma 1. *Suppose that Assumption* (A) *is fulfilled. Then the stopping time given by* (4.1) *belongs to the class* $\boldsymbol{T}$, *that is, it follows that for any* $\theta \in \Theta$,

$$\frac{\tau(u)}{u} \to \frac{1}{\alpha\dot{h}(\theta)} > 0 \textit{ in probability as } u \to \infty \textit{ under } P_\theta,$$

and

$$\frac{E_\theta(\tau(u))}{u} \to \frac{1}{\alpha\dot{h}(\theta)} > 0 \textit{ as } u \to \infty.$$

Next theorem shows that the large deviation probability of the sequential maximum likelihood estimator decays exponentially fast as $u \to \infty$.

Theorem 2. *Suppose that* $P_\theta(\tau(u) < \infty) = 1$ *for any* $\theta \in \Theta$, *where the stopping time* $\tau(u)$ *is given by* (4.1). *If Assumption* (A) *is fulfilled then for all sufficiently small* $\varepsilon > 0$, *we have*

$$\lim_{u\to\infty}\frac{1}{E_\theta(\tau(u))}\log P_{\theta,\tau(u)}(\left|\hat{\theta}_{\tau(u)}-\theta\right|>\varepsilon)=-B(\theta,\varepsilon)$$

Proof. From lemma 1, the stopping time $\tau(u)$ given by (4.1) belongs to the class $\boldsymbol{T}$ and the sequence of the maximum likelihood estimator $\{\hat{\theta}_{\tau(u)}:u\in\Gamma\}$ is consistent for θ. Therefore, by Theorem 1 it is sufficient to show that

$$\limsup_{u\to\infty}\frac{1}{E_\theta(\tau(u))}\log P_{\theta,\tau(u)}(\left|\hat{\theta}_{\tau(u)}-\theta\right|>\varepsilon)\le -B(\theta,\varepsilon).$$

It follows that

$$\begin{aligned}P_{\theta,\tau(u)}(\left|\hat{\theta}_{\tau(u)}-\theta\right|>\varepsilon)&=P_{\theta,\tau(u)}(\hat{\theta}_{\tau(u)}>\theta+\varepsilon)+P_{\theta,\tau(u)}(\hat{\theta}_{\tau(u)}<\theta-\varepsilon)\\&=P_{\theta,\tau(u)}(\dot{l}_{\tau(u)}(\theta+\varepsilon)>0)+P_{\theta,\tau(u)}(\dot{l}_{\tau(u)}(\theta-\varepsilon)<0)\\&=I_1+I_2,\end{aligned}\tag{4.8}$$

where

$$I_1=P_{\theta,\tau(u)}(\dot{l}_{\tau(u)}(\theta+\varepsilon)>0)$$

and

$$I_2=P_{\theta,\tau(u)}(\dot{l}_{\tau(u)}(\theta-\varepsilon)<0).$$

By Markov inequality, we obtain

$$\begin{aligned}I_1&\le\inf_{s>0}E_\theta(\exp(s\dot{l}_{\tau(u)}(\theta+\varepsilon)))\\&=\inf_{s>0}\phi_{\dot{l}_{\tau(u)}(\theta+\varepsilon),\theta}(s).\end{aligned}\tag{4.9}$$

According to (4.2) and (4.6), it follows that

$$\begin{aligned}\phi_{\dot{l}_{\tau(u)}(\theta+\varepsilon),\theta}(s)&=E_\theta[\exp\left(s(X(\tau(u))-\dot{f}(\theta+\varepsilon)\tau(u))\right)]\\&=E_{\mu_{\tau(u)}}[\exp\left((s+\theta)X(\tau(u))-(f(\theta)+s\dot{f}(\theta+\varepsilon))\tau(u)\right)]\\&=E_{\mu_{\tau(u)}}[\exp\left(\alpha^{-1}(s+\theta)(D_{\tau(u)}+u)-(f(\theta)+s\dot{f}(\theta+\varepsilon)+(s+\theta)\alpha^{-1}\beta)\tau(u)\right)]\\&=E_{\mu_{\tau(u)}}[\exp\left((s+\theta)\alpha^{-1}(D_{\tau(u)}+u)-(h(\theta)+s\dot{h}(\theta+\varepsilon))\tau(u)\right)]\\&=\exp\left(\alpha^{-1}u(s+\theta-h^{-1}(h(\theta)+s\dot{h}(\theta+\varepsilon)))\right)\\&\quad\cdot E_{\mu_{\tau(u)}}[\exp\left(\alpha^{-1}(s+\theta-h^{-1}(h(\theta)+s\dot{h}(\theta+\varepsilon)))D_{\tau(u)}\right)\cdot g(X(\tau(u)),\tau(u))\\&\quad\cdot\exp\left(\alpha^{-1}h^{-1}(h(\theta)+s\dot{h}(\theta+\varepsilon))(D_{\tau(u)}+u)-(h(\theta)+s\dot{h}(\theta+\varepsilon))\tau(u)\right)]\\&=\exp\left(\alpha^{-1}u(s+\theta-\vartheta)\right)E_\vartheta[\exp\left(\alpha^{-1}(s+\theta-\vartheta)D_{\tau(u)}\right)]\\&=\exp\left(\alpha^{-1}u(s+\theta-\vartheta)\right)\phi_{D_{\tau(u)},\vartheta}(\alpha^{-1}(s+\theta-\vartheta)),\end{aligned}\tag{4.10}$$

where $\vartheta = h^{-1}(h(\theta) + s\dot{h}(\theta+\varepsilon))$.

Let $\psi_{\theta,\varepsilon}(s) = \alpha^{-1}(s+\theta-\vartheta)$. Since h is invertible and differentiable on Θ, it follows that

$$(\partial/\partial s)\,\psi_{\theta,\varepsilon}(s) = \alpha^{-1}u\left(1-\dot{h}(\theta+\varepsilon)/\dot{h}(h^{-1}(s\dot{h}(\theta+\varepsilon)+h(\theta)))\right).$$

It is easily seen that the equation $(\partial/\partial s)\,\psi_{\theta,\varepsilon}(s) = 0$ has a unique solution $s_0 = (h(\theta+\varepsilon)-h(\theta))/\dot{h}(\theta+\varepsilon) > 0$ and $\psi_{\theta,\varepsilon}(s)$ attains its minimum value for $s = s_0$. Hence,

$$\begin{aligned}
\inf_{s>0} \psi_{\theta,\varepsilon}(s) &= \psi_{\theta,\varepsilon}(s_0)\\
&= \alpha^{-1}u(s_0+\theta-\vartheta)\\
&= (\alpha^{-1}u(h(\theta+\varepsilon)-h(\theta)-s\dot{h}(\theta+\varepsilon)))/\dot{h}(\theta+\varepsilon)\\
&= -uc(\theta+\varepsilon)K(\theta+\varepsilon,\theta). \qquad (4.11)
\end{aligned}$$

By (4.9), (4.10), and (4.11), we have

$$\begin{aligned}
\log I_1 &\le \inf_{s>0} \log \phi_{l_{\tau(u)}(\theta+\varepsilon),\theta}(s)\\
&= \inf_{s>0}\left(\psi_{\theta,\varepsilon}(s) + \log \phi_{D_{\tau(u)},\vartheta}(\alpha^{-1}(s+\theta-\vartheta))\right)\\
&\le \psi_{\theta,\varepsilon}(s_0) + \log \phi_{D_{\tau(u)},\vartheta}(\alpha^{-1}(s_0+\theta-\vartheta)).\\
&\le -uc(\theta+\varepsilon)K(\theta+\varepsilon,\theta) + \log E_\vartheta\left(\exp(\left|\alpha^{-1}(s_0+\theta-\vartheta)\right| M_\theta(u)\right). \qquad (4.12)
\end{aligned}$$

Since the distribution of $M_\theta(u)$ under P_ϑ is independent of u and the stopping time $\tau(u)$ given by (4.1) belongs to the class $\boldsymbol{T}$, it follows that

$$\limsup_{u\to\infty} \frac{1}{E_\theta(\tau(u))} \log I_1 \le -\tilde{K}(\theta+\varepsilon,\theta). \qquad (4.13)$$

In a similar fashion, we have

$$\limsup_{u\to\infty} \frac{1}{E_\theta(\tau(u))} \log I_2 \le -\tilde{K}(\theta-\varepsilon,\theta). \qquad (4.14)$$

Hence, by (4.8), (4.13), and (4.14), we have

$$\lim_{u\to\infty} \frac{1}{E_\theta(\tau(u))} \log P_{\theta,\tau(u)}(\left|\hat{\theta}_{\tau(u)} - \theta\right| > \varepsilon) \le -B(\theta,\varepsilon).$$

This completes the proof. □

Now, let $I_u(\theta)$ be the Fisher information, that is, $I_u(\theta) := E_\theta\left((\frac{\partial}{\partial\theta}\log L_{\tau(u)}(\theta))^2\right)$. We have

$$I_u(\theta) = -E_\theta\left(\frac{\partial^2}{\partial\theta^2}\log L_{\tau(u)}(\theta)\right)$$

$$= \ddot{f}(\theta) E_\theta(\tau(u)).$$

We define $I(\theta) := \ddot{f}(\theta)$. It is easily seen that

$$K(\theta_1, \theta) = \frac{1}{2}(\theta_1 - \theta)^2 I(\theta)[1 + o(1)] \quad \text{as} \quad \theta_1 \to \theta. \tag{4.15}$$

Let $\varphi_{\tau(u)} = \varphi(\tau(u), X(\tau(u)))$ be an estimator of θ and let $\lambda_u = \lambda_u(\varepsilon, \theta)$ be defined by

$$P_\theta(|\varphi_{\tau(u)} - \theta| > \varepsilon) = P(|N(0,1)| > \varepsilon/\lambda_u), \tag{4.16}$$

where $N(0,1)$ is a normal random variable with mean 0 and unit variance. Following Bahadur we call λ_u *the effective standard deviation* of $\varphi_{\tau(u)}$. According to (4.16), it is clear that $\varphi_{\tau(u)}$ is consistent for θ if and only if $\lambda_u \to 0$ as $u \to \infty$. From the fact that for $x > 0$,

$$(1/x - 1/x^3)(2\pi)^{-1/2} \exp(-x^2/2) < P(|N(0,1)| > x) < (1/x)(2\pi)^{-1/2} \exp(-x^2/2),$$

if $\varphi_{\tau(u)}$ is consistent for θ then

$$\log P_\theta(|\varphi_{\tau(u)} - \theta| > \varepsilon) = -\frac{1}{2}\frac{\varepsilon^2}{\lambda_u^2}(1 + o(1)) \quad \text{as } u \to \infty. \tag{4.17}$$

By Theorem 1 and (4.15), if $\varphi_{\tau(u)}$ is consistent for θ with respect to any sequence of the stopping times satisfying $\{\tau(u) : u \in \Gamma\} \in \boldsymbol{T}$ then

$$\liminf_{\varepsilon \to 0} \liminf_{u \to \infty} \frac{1}{\varepsilon^2 E_\theta(\tau(u))} \log P_\theta(|\varphi_{\tau(u)} - \theta| > \varepsilon) \ge -\limsup_{\varepsilon \to 0} \frac{B(\theta, \varepsilon)}{\varepsilon^2}$$
$$= -\frac{1}{2} I(\theta).$$

Therefore, from (4.17) we have

$$\liminf_{\varepsilon \to 0} \liminf_{u \to \infty} \{E_\theta(\tau(u)) \lambda_u^2\} \ge I(\theta)^{-1}. \tag{4.18}$$

Inequality (4.18) gives us an asymptotic lower bound of the effective standard deviation of $\varphi_{\tau(u)}$. We shall say that a consistent estimator $\varphi_{\tau(u)}$ is *asymptotically efficient in the Bahadur sense* if

$$\lim_{\varepsilon \to 0} \lim_{u \to \infty} \{E_\theta(\tau(u)) \lambda_u^2\} = I(\theta)^{-1}$$

holds for any $\theta \in \Theta$.

By Theorem 2, we obtain the next theorem.

Theorem 3. *Suppose that Assumption* (A) *is fulfilled. Then the sequential maximum likelihood estimator* $\hat{\theta}_{\tau(u)}$ *is asymptotically efficient in the Bahadur sense among all estimators which belong to* $\boldsymbol{C}$, *where the stopping time* $\tau(u)$ *is given by* (4.1).

5. Examples

To illustrate our results, we give two examples.

Example 1. Let $X(t)$ be a Wiener process with drift θ and unit variance. Of course, this process belongs to the exponential class and the density function is given by

$$f(x,t,\theta)=\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{X(t)^2}{2t}\right)\exp\left(\theta X(t)-\frac{1}{2}\theta^2 t\right), \quad t\in T=[0,\infty),$$

where θ takes its value in a parameter space $\Theta\subset R^1$. We suppose that Θ is a subset of the set $\{\theta.\alpha\theta+\beta>0\}$. Since $X(t)$ is continuous with probability one, we have $P_\theta(D_{\tau(u)}=0)=1$ for any $\theta\in\Theta$. Therefore, it is easily seen that Assumption (A) is fulfilled and the sequential maximum likelihood estimator $\hat{\theta}_{\tau(u)}=X(\tau(u))/\tau(u)$ is asymptotically efficient in the Bahadur sense.

By the way, in this example we can directly derive the asymptotic behavior of tail probability of $\hat{\theta}_\tau$ as follows . It is known that the stopping time τ is distributed as the generalized inverse Gaussian distribution $N^-(-\frac{1}{2},u^2\alpha^{-2},(\theta+\alpha^{-1}\beta)^2)$(see Sørensen(1986)). We denote $N^-(\lambda,\chi,\psi)$ as the generalized inverse Gaussian distribution. This distribution has the density function

$$f(x;\lambda,\chi,\psi)=\frac{(\psi/\chi)^{\frac{\lambda}{2}}}{2K_\lambda(\sqrt{\chi\psi})}x^{\lambda-1}\exp(-\tfrac{1}{2}(\frac{\chi}{x}+\psi x))\cdot I_{(0,\infty)}(x), \qquad (5.1)$$

where K_λ is the modified Bessel function of the third kind. For details, see Jørgensen(1982). Since $\alpha X(\tau)+\beta\tau=u$ and $\hat{\theta}_{\tau(u)}=X(\tau(u))/\tau(u)$ we have $\hat{\theta}_\tau=\dfrac{\alpha^{-1}u}{\tau}-\alpha^{-1}\beta$. Without loss of generality, we assume that $\alpha>0$ in the following discussions, because if $\alpha<0$ then we can replace $I_{(0,\infty)}$ with $I_{(-\infty,0)}$ in (5.3) below. Under this assumption, it follows that $\theta+\alpha^{-1}\beta>0$. By property of the generalized inverse Gaussian distribution we have

$$\hat{\theta}_\tau+\alpha^{-1}\beta\sim N^-(\tfrac{1}{2},u\alpha^{-1}(\theta+\alpha^{-1}\beta)^2,u\alpha^{-1}). \qquad (5.2)$$

From (5.1), substitutions $\lambda = \frac{1}{2}$, $\chi = u\alpha^{-1}(\theta+\alpha^{-1}\beta)^2$, $\psi = u\alpha^{-1}$ give

$$f(x;\tfrac{1}{2},u\alpha^{-1}(\theta+\alpha^{-1}\beta)^2,u\alpha^{-1}) = \frac{(\theta+\alpha^{-1}\beta)^{-\frac{1}{2}}}{2K_{\frac{1}{2}}(u\alpha^{-1}(\theta+\alpha^{-1}\beta))} x^{-\frac{1}{2}} \exp(-\tfrac{1}{2}(\frac{u\alpha^{-1}(\theta+\alpha^{-1}\beta)^2}{x}+u\alpha^{-1}x)) \cdot I_{(0,\infty)}(x)$$

Using the fact that $K_{\frac{1}{2}}(x) = \sqrt{\frac{\pi}{2}}\, x^{-\frac{1}{2}} e^{-x}$ (see, Jørgensen(1986), pp. 170), we have

$$f(x;\tfrac{1}{2},u\alpha^{-1}(\theta+\alpha^{-1}\beta)^2,u\alpha^{-1})$$

$$= \sqrt{\frac{u\alpha^{-1}}{2\pi}}\, x^{-\frac{1}{2}} \exp(-\frac{u\alpha^{-1}(x-(\theta+\alpha^{-1}\beta))^2}{2x}) \cdot I_{(0,\infty)}(x)$$

$$= \frac{\sqrt{n}}{\sqrt{2\pi}} x^{-\frac{1}{2}} \exp(-\frac{n(x-m)^2}{2x}) \cdot I_{(0,\infty)}(x), \qquad (5.3)$$

where $m = \theta+\alpha^{-1}\beta$, $n = u\alpha^{-1}$. Shuster(1968) obtained the distribution function of inverse Gaussian distribution. We can derive the distribution function of the generalized inverse Gaussian distribution given by (5.3) along the lines of Shuster(1968) as follows. Let X be a random variable which is distributed as the generalized inverse Gaussian distribution given by (5.3). Let $F(c;m,n) = P(X < c)$. Note that

$$F(c;m,n) = F(\frac{c}{m};\, 1,\, mn\,). \qquad (5.4)$$

First we deal with the case m=1.

Case (i). Let $c \le 1$. Using substitution $y = \sqrt{x}$, we have

$$F(c;1,n) = \int_0^c \sqrt{\frac{n}{2\pi x}} \exp[-\frac{n(x-1)^2}{2x}]dx$$

$$= \int_0^{\sqrt{c}} \sqrt{\frac{2n}{\pi}} \exp[-\frac{n(y^2-1)^2}{2y^2}]dy.$$

Put $z = \dfrac{n(y^2-1)^2}{y^2}$ for $y \le \sqrt{c}$. Since $y \le \sqrt{c} \le 1$, we have

$$y^2 = \frac{2n+z-\sqrt{z^2+4nz}}{2n}, \quad \frac{y^2}{y^2+1}$$

$$= \frac{1}{2}\left(1+\sqrt{\frac{z}{z+4n}}\right),$$

and $\dfrac{n(c-1)^2}{c} < z < \infty$.

Hence, we have

$$F(c;1,n) = \int_{\frac{n(c-1)^2}{c}}^{\infty} \frac{1}{\sqrt{2\pi z}} \left(\frac{y^2}{y^2+1}\right) \exp(-\frac{z}{2})dz$$

$$= \int_{\frac{n(c-1)^2}{c}}^{\infty} \frac{1}{\sqrt{2\pi z}} \cdot \frac{1}{2}\left(1+\sqrt{\frac{z}{z+4n}}\right) \exp(-\frac{z}{2})dz$$

$$= \frac{1}{2}G(d) + \frac{1}{2}\exp(2n)G(d+4n), \tag{5.5}$$

where $G(d) = \int_d^{\infty} (2\pi t)^{-\frac{1}{2}} \exp(-\frac{t}{2})dt$ and $d = \dfrac{n(c-1)^2}{c}$.

Case(ii). Let $c > 1$. Since $d = \dfrac{n(\frac{1}{c}-1)^2}{\frac{1}{c}}$, by (5.5) we have

$$F(c;1,n) = F(1/c;1,n) + (\ F(c;1,n) - F(1/c;1,n)\ )$$

$$= \frac{1}{2}G(d) + \frac{1}{2}\exp(2n)G(d+4n) + P(\frac{1}{c} < X < c). \tag{5.6}$$

It is easy to see that $\frac{1}{c} < X < c$ if and only if $\dfrac{n(X-1)^2}{X} < d$. From the fact that the random variable $\frac{n(X-1)^2}{X}$ is distributed as a chi-squared distribution with one degree of freedom, it follows that

$$P(\frac{1}{c} < X < c) = P(\chi_1^2 < d) = 1 - G(d).$$

Hence, by (5.6) we obtain

$$F(c;1,n) = 1 - \frac{1}{2}G(d) + \frac{1}{2}\exp(2n)G(d+4n). \tag{5.7}$$

From (5.4), (5.5), and (5.7), it is immediately evident that the following results hold for any m>0.

$$F(c;m,n) = \frac{1}{2}G(d') + \frac{1}{2}\exp(2mn)G(d'+4mn), \quad 0 \le c \le m,$$

$$= 1 - \frac{1}{2}G(d') + \frac{1}{2}\exp(2mn)G(d'+4mn), \quad m < c < \infty. \tag{5.8}$$

where $d' = \frac{n(c-m)^2}{c}$. Since $P(\chi_1^2 > x) = 2P(N(0,1) > \sqrt{x}) = 2P(N(0,1) < -\sqrt{x})$ for x>0, by (5.8) it follows that for any c>0, m>0, n>0,

$$F(c;m,n) = \Phi\left(\sqrt{\frac{n}{c}}(c-m)\right) + \exp(2mn)\Phi\left(-\sqrt{\frac{n}{c}}(c+m)\right), \tag{5.9}$$

where Φ is the standard normal distribution function.

Now, we have

$$P_{\theta,\tau}(|\hat{\theta}_\tau - \theta| > \varepsilon) = P_{\theta,\tau}(\hat{\theta}_\tau + \alpha^{-1}\beta > \theta + \varepsilon + \alpha^{-1}\beta) + P_{\theta,\tau}(\hat{\theta}_\tau + \alpha^{-1}\beta < \theta - \varepsilon + \alpha^{-1}\beta)$$

$$= I_1 + I_2, \tag{5.10}$$

where $I_1 = P_{\theta,\tau}(\hat{\theta}_\tau + \alpha^{-1}\beta > \theta + \varepsilon + \alpha^{-1}\beta)$ and $I_2 = P_{\theta,\tau}(\hat{\theta}_\tau + \alpha^{-1}\beta < \theta - \varepsilon + \alpha^{-1}\beta)$.

From (5.2) and (5.9), we have

$$I_1 = 1 - \Phi\left(\varepsilon\sqrt{\frac{u\alpha^{-1}}{\theta+\varepsilon+\alpha^{-1}\beta}}\right)$$

$$-\exp\left(2u\alpha^{-1}(\theta+\alpha^{-1}\beta)\right)\Phi\left(-\sqrt{\frac{u\alpha^{-1}}{\theta+\varepsilon+\alpha^{-1}\beta}}\,(2(\theta+\alpha^{-1}\beta)+\varepsilon)\right).$$

From the fact that $(x^{-1}-x^{-3})\phi(x) < 1-\Phi(x) < x^{-1}\phi(x)$ for any x>0, where $\phi(x)$ is the density function of the standard normal distribution, we have

$$I_1 \le 1 - \Phi\left(\varepsilon\sqrt{\frac{u\alpha^{-1}}{\theta+\varepsilon+\alpha^{-1}\beta}}\right)$$

$$< \varepsilon^{-1}\left(\frac{u\alpha^{-1}}{\theta+\alpha^{-1}\beta+\varepsilon}\right)^{-\frac{1}{2}}\frac{1}{\sqrt{2\pi}}e^{-\frac{u\alpha^{-1}\varepsilon^2}{2(\theta+\alpha^{-1}\beta+\varepsilon)}}$$

$$= k_1 u^{-\frac{1}{2}} e^{-\frac{u\alpha^{-1}\varepsilon^2}{2(\theta+\alpha^{-1}\beta+\varepsilon)}},$$

and

$$I_1 > \left[\varepsilon^{-1}\left(\frac{u\alpha^{-1}}{\theta+\alpha^{-1}\beta+\varepsilon}\right)^{-\frac{1}{2}} - \varepsilon^{-3}\left(\frac{u\alpha^{-1}}{\theta+\alpha^{-1}\beta+\varepsilon}\right)^{-\frac{3}{2}}\right]\frac{1}{\sqrt{2\pi}}e^{-\frac{u\alpha^{-1}\varepsilon^2}{2(\theta+\alpha^{-1}\beta+\varepsilon)}}$$

$$-e^{\left(2u\alpha^{-1}(\theta+\alpha^{-1}\beta)\right)}\left(\frac{u\alpha^{-1}(2(\theta+\alpha^{-1}\beta)+\varepsilon)^2}{\theta+\alpha^{-1}\beta+\varepsilon}\right)^{-\frac{1}{2}}\frac{1}{\sqrt{2\pi}}e^{-\frac{u\alpha^{-1}(2(\theta+\alpha^{-1}\beta)+\varepsilon)^2}{2(\theta+\alpha^{-1}\beta+\varepsilon)}}$$

$$= (u^{-\frac{1}{2}}k_2 - u^{-\frac{3}{2}}k_3)e^{-\frac{u\alpha^{-1}\varepsilon^2}{2(\theta+\alpha^{-1}\beta+\varepsilon)}},$$

where k_1, k_2, and k_3 are positive and independent of u.

Hence, we have

$$I_1 = e^{-\frac{u\alpha^{-1}\varepsilon^2}{2(\theta+\alpha^{-1}\beta+\varepsilon)}} \cdot u^{-\frac{1}{2}} \cdot O_e(1) = e^{-uK(\theta+\varepsilon,\theta)} \cdot u^{-\frac{1}{2}} \cdot O_e(1). \tag{5.11}$$

Similarly, we have

$$I_2 = e^{-\frac{u\alpha^{-1}\varepsilon^2}{2(\theta+\alpha^{-1}\beta-\varepsilon)}} \cdot u^{-\frac{1}{2}} \cdot O_e(1) = e^{-uK(\theta-\varepsilon,\theta)} \cdot u^{-\frac{1}{2}} \cdot O_e(1). \tag{5.12}$$

By (5.10), (5.11), and (5.12), we have

$$\lim_{u\to\infty} \frac{1}{E_\theta(\tau(u))} \log P_{\theta,\tau}(\left|\hat{\theta}_\tau - \theta\right| > \varepsilon) = -B(\theta, \varepsilon).$$

Example 2. Let $X(t)$ be a Poisson process with intensity $\vartheta > 0$. Let $\theta = \log \vartheta$ and suppose that Θ is a subset of the set $\{\theta.\, \alpha e^\theta + \beta > 0\}$. The density function is given by

$$f(x,t,\theta) = \frac{t^{X(t)}}{X(t)!}\exp\left(\theta X(t) - e^\theta t\right), \quad t \in T = \{1,2,\cdots\},$$

Since the Poisson process has jumps of size one, we have $0 \le D_{\tau(u)} \le |\alpha|$ with probability one. Therefore Assumption (A) is fulfilled and the sequential maximum likelihood estimator $\hat{\theta}_{\tau(u)} = \log\left(X(\tau(u))/\tau(u)\right)$ is asymptotically efficient in the Bahadur sense.

References

Bahadur, R.R. (1960), Asymptotic efficiency of tests and estimates, *Sankhyā*. **22**, 229-252.

Bahadur, R.R. (1967), Rates of convergence of estimates and test statistics, *Ann. Math. Statist.* **38**, 303-324.

Bsawa, I.V. and Prakasa Rao, B.L.S. (1980), *Statistical Inference for Stochastic Processes.* Academic Press.

Doob, J.L. (1953), *Stochastic Process*, Wiley.

Jørgensen, B. (1982), *Statistical Properties of the Generalized Inverse Gaussian Distribution*, Lecture Notes in Statistics, 9, Springer.

Shuster, J. (1968). On the inverse Gaussian distribution function, *Journal of the American Statistcal Association*, 63, 1514-1516.

Sørensen, M. (1986), On sequential maximum likelihood estimation for exponential families of stochastic processes, *Inter. Statist. Rev.* **54**, 191-210.

Winkler, W. and Franz, J. (1979), Sequential estimation problems for the exponential class of processes with independent increments, *Scand. J. Statist.* **6**, 129-139.

Winkler, W., Franz, J. and Küchler, I. (1982), Sequential statistical procedures for processes of the exponential class with independent increments, *Math. Operationsforsch. Statist., Ser. Statistics.* **13**, 105-119.