Metrika (2011) 74:349–360
DOI 10.1007/s00184-010-0306-8

Asymptotic properties of maximum likelihood estimators based on progressive Type-II censoring

Chien-Tai Lin · N. Balakrishnan

Received: 5 March 2008 / Published online: 26 March 2010

Abstract Hoadley (Ann Math Stat 42:1977–1991, 1971) studied the weak law of large numbers for independent and non-identically distributed random variables. Using that result along with the missing information principle, we establish the consistency and asymptotic normality of maximum likelihood estimators based on progressively Type-II censored samples.

Keywords Asymptotic theory · Consistency · Maximum likelihood · Progressive Type-II censoring · Missing information principle

1 Introduction

In a life-testing experiment, suppose n identical units are put on a life-test under the progressive censoring scheme $(R_1, R_2, \ldots, R_m)$, where $1 \le m \le n$ and $\lim_{n\to\infty} m/n = \tau$ is pre-fixed. Here, the censoring numbers R_i's, though fixed, as n tends to infinity, are assumed to be fixed in proportion which means that R_i/n tends to a proportion τ_i as n tends to infinity. Then, as $n \to \infty$, these proportions are such that $\sum_i \tau_i$ goes to $1 - \tau$, where τ is the overall proportion of observations and therefore $1 - \tau$ is the overall proportion of censoring. Let $f(x;\theta)$ and $F(x;\theta)$ denote,

N. Balakrishnan is a visiting professor at King Saud University, Saudi Arabia and National Central University, Taiwan.

C.-T. Lin (✉)
Department of Mathematics, Tamkang University, Tamsui 25137, Taiwan
e-mail: chien@mail.tku.edu.tw

N. Balakrishnan
Department of Mathematics and Statistics, McMaster University,
Hamilton, ON L8S 4K1, Canada
e-mail: bala@univmail.cis.mcmaster.ca

respectively, the probability density function and the distribution function of the life time of the unit, where $-\infty < x < \infty$ and θ is a single parameter of interest. Then, the likelihood function of the observed sample $x_{1:m:n}, \ldots, x_{m:m:n}$ is

$$L(\theta; x_{1:m:n}, \ldots, x_{m:m:n}) = C \prod_{i=1}^{m} f(x_{i:m:n}; \theta) \{1 - F(x_{i:m:n}; \theta)\}^{R_i}, \quad x_{1:m:n} < \cdots < x_{m:m:n}, \tag{1}$$

where $C = n(n - R_1 - 1) \cdots (n - R_1 - R_2 - \cdots - R_{m-1} - m + 1)$; see for example, Balakrishnan and Aggarwala (2000) and Balakrishnan (2007). Under a variety of conditions, the asymptotic properties of the maximum likelihood estimator of θ have been studied for different types of data including complete data and also Type-I, Type-II, multiply Type-II, random, and progressive censoring schemes; see for example, Cramér (1946), Halperin (1952), Hoadley (1971), Sen (1976), Bhattacharyya (1985), and Kong and Fei (1996), and the references contained therein. Based on martingales and functional central limit theorems (see Sen 1976), the derivation of these results in the context of random censoring and progressive censoring involves tedious manipulations as well as stronger regularity conditions than what are really necessary. Bhattacharyya (1985) used a result of Sethuraman (1961) concerning the conditional and joint distribution of random vectors to establish the asymptotic properties of maximum likelihood estimators (MLE) and some modified MLEs based on Type-II censored data. Unfortunately, these results do not extend readily to the case of progressive censoring. In this paper, we use the weak law of large numbers for independent and non-identically distributed (i.ni.d.) random variables derived by Hoadley (1971) along with the missing information principle to establish the consistency and asymptotic normality of the MLE of θ based on progressively Type-II censored samples. This result may be further generalized to the case of several unknown parameters by suitably modifying the conditions on which it has been established.

2 Consistency

The approach to consistency will follow the same lines as in Cramér (1946). The conditions are:

$A1$: For almost all x, the derivatives $\frac{\partial}{\partial\theta} \ln f(x;\theta)$, $\frac{\partial^2}{\partial\theta^2} \ln f(x;\theta)$, and $\frac{\partial^3}{\partial\theta^3} \ln f(x;\theta)$ all exist for every θ belonging to a non-degenerate interval I;

$A2$: For every θ in I, we have

$$\left|\frac{\partial}{\partial\theta} f(x;\theta)\right| \le G_1, \quad \left|\frac{\partial^2}{\partial\theta^2} f(x;\theta)\right| \le G_2, \quad \left|\frac{\partial^3}{\partial\theta^3} f(x;\theta)\right| \le G_3,$$

where

$$\int G_i(x) d\mu(x) < \infty \quad \text{for } i = 1, 2, 3,$$

and μ is taken to be Lebesgue measure;

$A3$: For every θ in I and positive constants δ and K, we have

$$\left|\frac{\partial}{\partial\theta}\ln f(x;\theta)\right| \leq G_1^*, \quad \left|\frac{\partial^2}{\partial\theta^2}\ln f(x;\theta)\right| \leq G_2^*, \quad \left|\frac{\partial^3}{\partial\theta^3}\ln f(x;\theta)\right| \leq G_3^*,$$

and

$$\int \left|G_i^*(x)\right|^{1+\delta} f(x;\theta)d\mu(x) \leq K \quad \text{for } i = 1, 2, 3;$$

$A4$: For every θ in I and positive constant M, $\frac{1}{1-F(x;\theta)}$ is bounded by $\eta(x)$, where

$$\int \eta(x) f(x;\theta)d\mu(x) \leq M;$$

$A5$: For every θ in I, the integral

$$\gamma^2 = \int \left[\frac{\partial}{\partial\theta}\ln f(x;\theta)\right]^2 f(x;\theta)d\mu(x)$$

is finite and positive.

In addition to these conditions, the following lemma and theorem are needed for establishing the consistency.

Lemma 1 *Assume conditions $A2 - A4$ are valid, and there exists a measurable and integrable function T with $\int T(x_{1:m:n}, \ldots, x_{m:m:n})d\mu(x) < \infty$, and Q is a positive constant independent of θ. Then, we have*

$$\frac{1}{n}\left|\frac{\partial^3}{\partial\theta^3}\ln L(\theta; x_{1:m:n}, \ldots, x_{m:m:n})\right| \leq T(x_{1:m:n}, \ldots, x_{m:m:n})$$

and $E[T(X_{1:m:n}, \ldots, X_{m:m:n})] \leq Q$.

Proof For simplicity, we replace $x_{i:m:n}$ by x_i throughout the proof. From the likelihood function in (1), we have

$$\frac{1}{n}\left|\frac{\partial^3}{\partial\theta^3}\ln L(\theta; x_1, \ldots, x_m)\right| \leq \frac{1}{n}\sum_{i=1}^{m}\left|\frac{\partial^3}{\partial\theta^3}\ln f(x_i, \theta)\right|$$
$$+\frac{1}{n}\sum_{i=1}^{m} R_i \left|\frac{\partial^3}{\partial\theta^3}\ln[1 - F(x_i, \theta)]\right|.$$

Under conditions $A3$ and $A4$, we have

$$\left|\frac{\partial^3}{\partial\theta^3}\ln f(x_i;\theta)\right| \leq G_3^*(x_i),$$

and

$$\begin{aligned}
&\left|\frac{\partial^3}{\partial\theta^3}\ln[1-F(x_i,\theta)]\right| \\
&= \left|\frac{\frac{\partial^3}{\partial\theta^3}[1-F(x_i,\theta)]}{1-F(x_i,\theta)} - 3\frac{\frac{\partial}{\partial\theta}[1-F(x_i,\theta)]\frac{\partial^2}{\partial\theta^2}[1-F(x_i,\theta)]}{[1-F(x_i,\theta)]^2} + 2\frac{\left(\frac{\partial}{\partial\theta}[1-F(x_i,\theta)]\right)^3}{[1-F(x_i,\theta)]^3}\right| \\
&\le \left|\frac{\partial^3}{\partial\theta^3}F(x_i,\theta)\right|\eta(x_i) + 3\left|\frac{\partial}{\partial\theta}F(x_i,\theta)\right|\left|\frac{\partial^2}{\partial\theta^2}F(x_i,\theta)\right|\eta^2(x_i) + 2\left|\frac{\partial}{\partial\theta}F(x_i,\theta)\right|^3\eta^3(x_i) \\
&= \left|\frac{\partial^3}{\partial\theta^3}\int\limits_{A_i} f(x,\theta)d\mu(x)\right|\eta(x_i)+3\left|\frac{\partial}{\partial\theta}\int\limits_{A_i} f(x,\theta)d\mu(x)\right|\left|\frac{\partial^2}{\partial\theta^2}\int\limits_{A_i} f(x,\theta)d\mu(x)\right|\eta^2(x_i) \\
&\quad +2\left|\frac{\partial}{\partial\theta}\int\limits_{A_i} f(x,\theta)d\mu(x)\right|^3\eta^3(x_i), \qquad (2)
\end{aligned}$$

where $A_i = \{x : -\infty < x \le x_i\}$ for $i = 1, \ldots, m$.

By Lebesgue dominated convergence theorem, condition $A2$ ensures that

$$\int \frac{\partial^i}{\partial\theta^i} f(x,\theta)d\mu(x) = \frac{\partial^i}{\partial\theta^i}\int f(x,\theta)d\mu(x) \quad \text{for } i = 1, 2, 3. \qquad (3)$$

Therefore, the expression on the RHS of Eq. (2) can be further simplified as

$$\begin{aligned}
&\left|\int\limits_{A_i} \frac{\partial^3}{\partial\theta^3} f(x,\theta)d\mu(x)\right|\eta(x_i) + 3\left|\int\limits_{A_i} \frac{\partial}{\partial\theta} f(x,\theta)d\mu(x)\right|\left|\int\limits_{A_i} \frac{\partial^2}{\partial\theta^2} f(x,\theta)d\mu(x)\right|\eta^2(x_i) \\
&\quad + 2\left|\int\limits_{A_i} \frac{\partial}{\partial\theta} f(x,\theta)d\mu(x)\right|^3\eta^3(x_i) \\
&\le \left|\int G_3(x)d\mu(x)\right|\eta(x_i) + 3\left|\int G_1(x)d\mu(x)\right|\left|\int G_2(x)d\mu(x)\right|\eta^2(x_i) \\
&\quad + 2\left|\int G_1(x)d\mu(x)\right|^3\eta^3(x_i) \equiv \nu(x_i).
\end{aligned}$$

Let

$$T(x_1,\ldots,x_m) \equiv \frac{1}{n}\sum_{i=1}^{m}\left[G_3^*(x_i) + R_i\nu(x_i)\right].$$

It is clear that $\int \nu(x) f(x;\theta) d\mu(x) \leq M^*$, where M^* is a positive constant independent of θ, and thus the lemma follows. □

The next theorem is the result of Hoadley (1971) which is a weak law of large numbers for i.ni.d. random variables.

Theorem 1 (Theorem A.5 of Hoadley 1971) *Let $\{Y_k : k = 1, 2, \ldots\}$ be independent random variables, which are defined on the probability space $(\Omega, \mathcal{F}, P_\theta)$, and take on values in a measure space $(\mathcal{Y}, \mathcal{A}, \mu)$. Let $H_k : \mathcal{Y} \times S \to \mathcal{R}^1$, where $S \subset \mathcal{R}^p$ is compact, and let $h_k(\mathbf{s}) = E H_k(Y_k, \mathbf{s})$. Assume:*

(a) *For each $\mathbf{s} \in S$, $H_k(\cdot, \mathbf{s})$ is $\mathcal{A}$ measurable;*
(b) *$H_k(Y_k, \cdot)$ is continuous on S, uniformly in k, a.s. [P];*
(c) *There exist measurable $B_k : \mathcal{Y} \to \mathcal{R}^1$ for which $|H_k(\cdot, \mathbf{s})| < B_k(\cdot)$ for all $\mathbf{s} \in S$ and $E|B_k(Y_k)|^{1+\delta} \leq K$, where K and δ are positive constant.*

Then:

(i) *$h_k(\cdot)$ is continuous on S, uniformly in k;*
(ii)

$$\sup\left\{\left|\frac{\sum_{k=1}^n H_k(Y_k, \mathbf{s})}{n} - \frac{\sum_{k=1}^n h_k(\mathbf{s})}{n}\right| : \mathbf{s} \in S\right\} \xrightarrow{p} 0,$$

where the notation $\xrightarrow{p}$ denotes convergence in probability.

We are now ready to prove the following theorem. Note that we shall only consider the case $1 \leq m < n$ in the proof since $m = n$ is simply the complete sample case.

Theorem 2 *If conditions $A1 - A5$ are satisfied, then the likelihood equation*

$$\frac{\partial}{\partial\theta} \ln L(\theta; x_{1:m:n}, \ldots, x_{m:m:n}) = \sum_{i=1}^{m}\left\{\frac{\partial}{\partial\theta} \ln f(x_{i:m:n}, \theta) + R_i \frac{\partial}{\partial\theta} \ln[1 - F(x_{i:m:n}, \theta)]\right\} = 0 \tag{4}$$

has a solution, $\widehat{\theta}_n$, which converges in probability to the true value of the parameter θ, say θ_0.

Proof From the Taylor expansion and Lemma 1, we can write

$$\begin{aligned}\frac{1}{n}\frac{\partial}{\partial\theta} \ln L(\theta; x_{1:m:n}, \ldots, x_{m:m:n}) &= \frac{1}{n}\frac{\partial}{\partial\theta} \ln L(\theta; x_{1:m:n}, \ldots, x_{m:m:n})\Big|_{\theta=\theta_0} \\ &\quad + \frac{(\theta-\theta_0)}{n}\frac{\partial^2}{\partial\theta^2} \ln L(\theta; x_{1:m:n}, \ldots, x_{m:m:n})\Big|_{\theta=\theta_0} \\ &\quad + \frac{1}{2}\Delta(\theta-\theta_0)^2 T(x_{1:m:n}, \ldots, x_{m:m:n}) \\ &= B_0 + B_1(\theta-\theta_0) + \frac{1}{2}\Delta(\theta-\theta_0)^2 B_2,\end{aligned} \tag{5}$$

where $|\Delta| < 1$, and

$$B_0 = \frac{1}{n}\sum_{i=1}^{m}\left\{\frac{\partial}{\partial\theta}\ln f(x_{i:m:n};\theta) + R_i\frac{\partial}{\partial\theta}\ln[1 - F(x_{i:m:n};\theta)]\right\}\Bigg|_{\theta=\theta_0},$$

$$B_1 = \frac{1}{n}\sum_{i=1}^{m}\left\{\frac{\partial^2}{\partial\theta^2}\ln f(x_{i:m:n};\theta) + R_i\frac{\partial^2}{\partial\theta^2}\ln[1 - F(x_{i:m:n};\theta)]\right\}\Bigg|_{\theta=\theta_0},$$

$$B_2 = T(x_{1:m:n},\ldots,x_{m:m:n}).$$

We shall prove, in probability, that $B_0 \xrightarrow{p} 0$, $B_1 \xrightarrow{p} -\zeta_1^2$ and $B_2 \xrightarrow{p} \zeta_2$, where ζ_1 and ζ_2 are constants that will be defined later.

Given a progressively Type-II censored sample $x_{1:m:n},\ldots,x_{m:m:n}$, we use the missing information principle (see e.g., Louis 1982; Tanner 1993; Ng et al. 2002) to write the observed information through the fact

$$\sum_{i=1}^{m}\ln f(x_{i:m:n};\theta) = \sum_{i=1}^{n}\ln f(w_i;\theta) - \sum_{i=1}^{m}\sum_{j=1}^{R_i}\ln f(y_{ij};\theta|X_{i:m:n} = x_{i:m:n}), \quad (6)$$

where $w_1,\ldots,w_n$ can be considered as a complete random sample of size n from $F(x;\theta)$ and $y_{i1},\ldots,y_{iR_i}$ can be considered as a complete random sample of size R_i, $i = 1,\ldots,m$, from the left-truncated population with density function

$$\psi_i(y;\theta) = \frac{f(y;\theta)}{1 - F(x_{i:m:n};\theta)}, \quad y > x_{i:m:n}.$$

Moreover, the sequences of random variables W's and Y's are independent. For simplicity in notation, we use θ and $f(y_{ij};\theta|x_{i:m:n})$, $i = 1,\ldots,m$ and $j = 1,\ldots,R_i$, instead of θ_0 and $f(y_{ij};\theta|X_{i:m:n} = x_{i:m:n})$ in what follows.

Thus, B_0 can be reexpressed as

$$\begin{aligned} B_0 = \frac{1}{n}\Bigg\{&\sum_{i=1}^{n}\frac{\partial}{\partial\theta}\ln f(w_i;\theta) - \sum_{i=1}^{m}\sum_{j=1}^{R_i}\frac{\partial}{\partial\theta}\ln f(y_{ij};\theta|x_{i:m:n}) \\ &+ \sum_{i=1}^{m}R_i\frac{\partial}{\partial\theta}\ln[1 - F(x_{i:m:n};\theta)]\Bigg\} \\ \equiv \frac{1}{n}&(B_{01} - B_{02}), \end{aligned}$$

where

$$B_{01} = \sum_{i=1}^{n}\frac{\partial}{\partial\theta}\ln f(w_i;\theta)$$

and

$$B_{02} = \sum_{i=1}^{m}\sum_{j=1}^{R_i} \frac{\partial}{\partial\theta} \ln f(y_{ij};\theta|x_{i:m:n}) - \sum_{i=1}^{m} R_i \frac{\partial}{\partial\theta} \ln[1 - F(x_{i:m:n};\theta)].$$

From the result of Cramér (1946), we have $B_{01}/n \xrightarrow{p} 0$. It now suffices to show that $B_{02}/n \xrightarrow{p} 0$. Since we can write B_{02} as

$$\begin{aligned} B_{02} &= \sum_{i=1}^{m}\sum_{j=1}^{R_i} \frac{\partial}{\partial\theta} \ln f(y_{ij};\theta|x_{i:m:n}) - \sum_{i=1}^{m} R_i E\left[\frac{\partial}{\partial\theta} \ln f(Y_{i1};\theta|x_{i:m:n})\right] \\ &\quad + \sum_{i=1}^{m} R_i E\left[\frac{\partial}{\partial\theta} \ln f(Y_{i1};\theta|x_{i:m:n})\right] - \sum_{i=1}^{m} R_i \frac{\partial}{\partial\theta} \ln[1 - F(x_{i:m:n};\theta)] \end{aligned}$$

and that Eq. (3) implies that

$$\begin{aligned} E\left[\frac{\partial}{\partial\theta} \ln f(Y_{i1};\theta|x_{i:m:n})\right] &= \int_{B_i} \frac{\partial}{\partial\theta} \ln f(y;\theta) \frac{f(y;\theta)}{1 - F(x_{i:m:n};\theta)} d\mu(y) \\ &= \frac{1}{1 - F(x_{i:m:n};\theta)} \int_{B_i} \frac{\partial}{\partial\theta} f(y;\theta) d\mu(y) \\ &= \frac{\frac{\partial}{\partial\theta} \int_{B_i} f(y;\theta) d\mu(y)}{1 - F(x_{i:m:n};\theta)} \\ &= \frac{\frac{\partial}{\partial\theta}[1 - F(x_{i:m:n};\theta)]}{1 - F(x_{i:m:n};\theta)} \\ &= \frac{\partial}{\partial\theta} \ln[1 - F(x_{i:m:n};\theta)], \end{aligned} \tag{7}$$

where $B_i = \{y : x_{i:m:n} < y < \infty\}$ for $i = 1, \ldots, m$, Theorem 1 implies that

$$\begin{aligned} \sup\left\{\left|\frac{B_{02}}{n-m}\right| : \theta \in I\right\} = \sup\Bigg\{\Bigg| &\frac{\sum_{i=1}^{m}\sum_{j=1}^{R_i} \frac{\partial}{\partial\theta} \ln f(y_{ij};\theta|x_{i:m:n})}{n-m} \\ &- \frac{\sum_{i=1}^{m} R_i E\left[\frac{\partial}{\partial\theta} \ln f(Y_{i1};\theta|x_{i:m:n})\right]}{n-m}\Bigg| : \theta \in I\Bigg\} \xrightarrow{p} 0; \end{aligned}$$

so, $B_{02}/n \xrightarrow{p} 0$. Consequently, by Slutsky's theorem, B_0 converges to zero in probability.

Similar arguments can be used to establish the convergence of B_1 and B_2. First, consider $B_1 = (B_{11} - B_{12})/n$, where

$$B_{11} = \sum_{i=1}^{n} \frac{\partial^2}{\partial \theta^2} \ln f(w_i; \theta)$$

and

$$B_{12} = \sum_{i=1}^{m} \left\{ \sum_{j=1}^{R_i} \frac{\partial^2}{\partial \theta^2} \ln f(y_{ij}; \theta | x_{i:m:n}) - R_i \frac{\partial^2}{\partial \theta^2} \ln[1 - F(x_{i:m:n}; \theta)] \right\}.$$

From the result of Cramér (1946) once again, we have

$$\frac{B_{11}}{n} \xrightarrow{p} E\left[\frac{\partial^2}{\partial \theta^2} \ln f(X; \theta)\right] = -\gamma^2.$$

The term B_{12}/n can be rewritten as

$$\frac{n-m}{n} \left\{ \frac{\sum_{i=1}^{m} \sum_{j=1}^{R_i} \frac{\partial^2}{\partial \theta^2} \ln f(y_{ij}; \theta | x_{i:m:n})}{n-m} - \frac{\sum_{i=1}^{m} R_i E\left[\frac{\partial^2}{\partial \theta^2} \ln f(Y_{i1}; \theta | x_{i:m:n})\right]}{n-m} \right\}$$
$$-\frac{1}{n} \sum_{i=1}^{m} R_i \left\{ \frac{\partial^2}{\partial \theta^2} \ln[1 - F(x_{i:m:n}; \theta)] - E\left[\frac{\partial^2}{\partial \theta^2} \ln f(Y_{i1}; \theta | x_{i:m:n})\right] \right\}. \quad (8)$$

It follows from Theorem 1 that the first term in (8) converges in probability to zero as $n \to \infty$. Equations (7) and (3), respectively, lead to

$$\frac{\partial^2}{\partial \theta^2} \ln[1 - F(x_{i:m:n}; \theta)] = \frac{\frac{\partial^2}{\partial \theta^2}[1 - F(x_{i:m:n}; \theta)]}{1 - F(x_{i:m:n}; \theta)} - \left\{ \frac{\partial}{\partial \theta} \ln[1 - F(x_{i:m:n}; \theta)] \right\}^2$$
$$= \frac{\frac{\partial^2}{\partial \theta^2}[1 - F(x_{i:m:n}; \theta)]}{1 - F(x_{i:m:n}; \theta)} - \left\{ E\left[\frac{\partial}{\partial \theta} \ln f(Y_{i1}; \theta | x_{i:m:n})\right] \right\}^2$$

and

$$E\left[\frac{\partial^2}{\partial \theta^2} \ln f(Y_{i1}; \theta | x_{i:m:n})\right]$$
$$= \int_{B_i} \left[\frac{\frac{\partial^2}{\partial \theta^2} f(y; \theta)}{f(y; \theta)} - \left(\frac{\partial}{\partial \theta} \ln f(y; \theta) \right)^2 \right] \frac{f(y; \theta)}{1 - F(x_{i:m:n}; \theta)} d\mu(y)$$
$$= \int_{B_i} \frac{\frac{\partial^2}{\partial \theta^2} f(y; \theta)}{f(y; \theta)} \frac{f(y; \theta)}{1 - F(x_{i:m:n}; \theta)} d\mu(y) - \int_{B_i} \left[\frac{\partial}{\partial \theta} \ln f(y; \theta) \right]^2 \frac{f(y; \theta)}{1 - F(x_{i:m:n}; \theta)} d\mu(y)$$
$$= \frac{\frac{\partial^2}{\partial \theta^2}[1 - F(x_{i:m:n}; \theta)]}{1 - F(x_{i:m:n}; \theta)} - \int_{B_i} \left[\frac{\partial}{\partial \theta} \ln f(y; \theta) \right]^2 \frac{f(y; \theta)}{1 - F(x_{i:m:n}; \theta)} d\mu(y),$$

and hence the difference of these two terms can be further simplified as

$$E\left[\frac{\partial}{\partial\theta}\ln f(Y_{i1};\theta|x_{i:m:n})\right]^2-\left\{E\left[\frac{\partial}{\partial\theta}\ln f(Y_{i1};\theta|x_{i:m:n})\right]\right\}^2$$
$$=Var\left[\frac{\partial}{\partial\theta}\ln f(Y_{i1};\theta|x_{i:m:n})\right],$$

which, from the conditions $A4$ and $A5$, is clearly bounded and independent of θ. Thus,

$$\frac{1}{n-m}\sum_{i=1}^{m}R_i\left\{\frac{\partial^2}{\partial\theta^2}\ln[1-F(x_{i:m:n};\theta)]-E\left[\frac{\partial^2}{\partial\theta^2}\ln f(Y_{i1};\theta|x_{i:m:n})\right]\right\}$$
$$=\frac{1}{n-m}\sum_{i=1}^{m}\sum_{j=1}^{R_i}Var\left[\frac{\partial}{\partial\theta}\ln f(Y_{ij};\theta|x_{i:m:n})\right] \tag{9}$$

converges to a bounded value, say Λ. Combining these results and setting $\gamma^2+(1-\tau)\Lambda\equiv\zeta_1^2$, we obtain $B_1\xrightarrow{p}-\zeta_1^2$. Finally, following the same lines as for the convergence of B_1, we have that B_2 converges in probability to a bounded value, say ζ_2. The argument given in Cramér (1946) can then be employed to show that (4) has a solution, $\widehat{\theta}_n$, which converges in probability to θ_0. □

It should be noted that Theorem 2 is established only for a solution of the likelihood equation, and this solution need not be the unique MLE (see e.g., Ferguson 1996).

3 Asymptotic normality

The approach to asymptotic normality of the MLE of θ is related to the result of Hoadley (1971) and Slutsky's theorem in the multivariate case (see e.g., Serfling 1980). Hoadley (1971) gave the Liapounov form of the multivariate central limit theorem which plays an important role in establishing that MLEs are asymptotically normal in the i.ni.d. case.

Theorem 3 (Theorem A.6 of Hoadley 1971) *Let* $\{\boldsymbol{X}_k : k=1,2,\ldots\}$ *be independent* p*-dimensional random vectors with* $E\boldsymbol{X}_k=\boldsymbol{0}$, $Cov(\boldsymbol{X}_k)=\boldsymbol{\Gamma}_k$. *Assume*:

(a) $\sum_{k=1}^{n}\overline{\boldsymbol{\Gamma}}_k/n\longrightarrow\overline{\boldsymbol{\Gamma}}$, *where* $\overline{\boldsymbol{\Gamma}}$ *is positive definite*;
(b) *For some* $\delta>0$, $\sum_k E|\boldsymbol{\lambda}'\boldsymbol{X}_k|^{2+\delta}/n^{(2+\delta)/2}\longrightarrow 0$ *for all* $\boldsymbol{\lambda}\in\mathcal{R}^p$.

Then, $\sum_k \boldsymbol{X}_k/\sqrt{n}\xrightarrow{d}N(\boldsymbol{0},\overline{\boldsymbol{\Gamma}})$, *where* $\xrightarrow{d}$ *denotes convergence in distribution.*

The following result allows the convergence of univariate distribution functions to be extendable to that of convergence of multivariate distribution functions.

Theorem 4 (Slutsky's Theorem) *Let* $(X_n, Y_n), n=1,2,\ldots,$ *and* (X,Y) *be two random vectors defined on a probability space. Suppose that* $X_n\xrightarrow{d}X$ *and* $Y_n\xrightarrow{d}Y$.

If X_n and Y_n are independent for each n, then

$$(X_n, Y_n) \xrightarrow{d} (X^*, Y^*),$$

where $X^* \sim X$, $Y^* \sim Y$, and X^* and Y^* are independent in the same space.

The above two results are needed in the following proof of the asymptotic normality of the MLE of θ.

Theorem 5 *Under regularity conditions $A1 - A5$, the likelihood equation in (4) has a solution, $\widehat{\theta}$, which is asymptotically normally distributed.*

Proof From (5), we have

$$\zeta_1\sqrt{n}(\widehat{\theta}-\theta_0) = \frac{\frac{\sqrt{n}}{\zeta_1}B_0}{-\frac{B_1}{\zeta_1^2}-\frac{\Delta}{2\zeta_1^2}B_2(\widehat{\theta}-\theta_0)} = \frac{\frac{1}{\zeta_1\sqrt{n}}\sum_{i=1}^m\left\{\frac{\partial}{\partial\theta}\ln f(x_{i:m:n};\theta)+R_i\frac{\partial}{\partial\theta}\ln[1-F(x_{i:m:n};\theta)]\right\}}{-\frac{B_1}{\zeta_1^2}-\frac{\Delta}{2\zeta_1^2}B_2(\widehat{\theta}-\theta_0)}, \tag{10}$$

where $\zeta_1^2=\gamma^2+(1-\tau)\Lambda$. The denominator of the right-hand side of (10) converges in probability to 1, so that we only need to show that the numerator is asymptotically normal with zero mean and unit variance.

From Eqs. (6) and (7), we have

$$\begin{aligned}
&\sum_{i=1}^m\left\{\frac{\partial}{\partial\theta}\ln f(x_{i:m:n};\theta)+R_i\frac{\partial}{\partial\theta}\ln[1-F(x_{i:m:n};\theta)]\right\}\\
&=\sum_{i=1}^n\frac{\partial}{\partial\theta}\ln f(w_i;\theta)-\sum_{i=1}^m\sum_{j=1}^{R_i}\frac{\partial}{\partial\theta}\ln f(y_{ij};\theta|x_{i:m:n})\\
&\quad+\sum_{i=1}^m\sum_{j=1}^{R_i}E\left[\frac{\partial}{\partial\theta}\ln f(Y_{ij};\theta|x_{i:m:n})\right]=\sum_{i=1}^n\frac{\partial}{\partial\theta}\ln f(w_i;\theta)\\
&\quad-\sum_{i=1}^m\sum_{j=1}^{R_i}\left\{\frac{\partial}{\partial\theta}\ln f(y_{ij};\theta|x_{i:m:n})-E\left[\frac{\partial}{\partial\theta}\ln f(Y_{ij};\theta|x_{i:m:n})\right]\right\}.
\end{aligned}$$

It follows from the result of Cramér (1946) that $\frac{1}{\sqrt{n}}\sum_{i=1}^n\frac{\partial}{\partial\theta}\ln f(w_i;\theta)$ is asymptotically normal with mean 0 and variance γ^2.

Conditions $A3$ and $A4$ imply that $E|\frac{\partial}{\partial\theta}\ln f(Y_{ij};\theta|x_{i:m:n}) - E[\frac{\partial}{\partial\theta}\ln f(Y_{ij};\theta|x_{i:m:n})]|^3$ is bounded and independent of θ, say K^*, which leads to

$$\frac{1}{n^{3/2}}\sum_{i=1}^{m}\sum_{j=1}^{R_i}E\left|\frac{\partial}{\partial\theta}\ln f(Y_{ij};\theta|x_{i:m:n})-E\left[\frac{\partial}{\partial\theta}\ln f(Y_{ij};\theta|x_{i:m:n})\right]\right|^3$$
$$=\frac{n-m}{n^{3/2}}K^*\longrightarrow 0.$$

Combining this result with Eq. (9) that

$$\frac{1}{n-m}\sum_{i=1}^{m}\sum_{j=1}^{R_i}Var\left[\frac{\partial}{\partial\theta}\ln f(Y_{ij};\theta|x_{i:m:n})\right]\longrightarrow\Lambda,$$

it then follows from Theorem 3 that

$$\frac{1}{\sqrt{n-m}}\sum_{i=1}^{m}\sum_{j=1}^{R_i}\left\{\frac{\partial}{\partial\theta}\ln f(y_{ij};\theta|x_{i:m:n})-E\left[\frac{\partial}{\partial\theta}\ln f(Y_{ij};\theta|x_{i:m:n})\right]\right\}\xrightarrow{d}N(0,\Lambda).$$

Applying Slutsky's theorem, we obtain

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{m}\sum_{j=1}^{R_i}\left\{\frac{\partial}{\partial\theta}\ln f(y_{ij};\theta|x_{i:m:n})-E\left[\frac{\partial}{\partial\theta}\ln f(Y_{ij};\theta|x_{i:m:n})\right]\right\}$$
$$=\frac{\sqrt{n-m}}{\sqrt{n}}\frac{1}{\sqrt{n-m}}\sum_{i=1}^{m}\sum_{j=1}^{R_i}\left\{\frac{\partial}{\partial\theta}\ln f(y_{ij};\theta|x_{i:m:n})-E\left[\frac{\partial}{\partial\theta}\ln f(Y_{ij};\theta|x_{i:m:n})\right]\right\}$$
$$\xrightarrow{d}N(0,(1-\tau)\Lambda).$$

Theorem 4 can now be applied to yield

$$\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{\partial}{\partial\theta}\ln f(w_i;\theta),\ \frac{1}{\sqrt{n}}\sum_{i=1}^{m}\sum_{j=1}^{R_i}\left\{\frac{\partial}{\partial\theta}\ln f(y_{ij};\theta|x_{i:m:n})\right.\right.$$
$$\left.\left.-\ E\left[\frac{\partial}{\partial\theta}\ln f(Y_{ij};\theta|x_{i:m:n})\right]\right\}\right)\xrightarrow{d}(X^*,Y^*),$$

where $X^*\sim N(0,\gamma^2)$, $Y^*\sim N(0,(1-\tau)\Lambda)$, and X^* and Y^* are independent. From the property of continuous mapping, we then have

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{\partial}{\partial\theta}\ln f(w_i;\theta)-\frac{1}{\sqrt{n}}\sum_{i=1}^{m}\sum_{j=1}^{R_i}\left\{\frac{\partial}{\partial\theta}\ln f(y_{ij};\theta|x_{i:m:n})\right.$$
$$\left.-E\left[\frac{\partial}{\partial\theta}\ln f(Y_{ij};\theta|x_{i:m:n})\right]\right\}\xrightarrow{d}X^*-Y^*.$$

Therefore,

$$\sum_{i=1}^{m}\left\{\frac{\partial}{\partial\theta}\ln f(x_{i:m:n};\theta)+R_i\frac{\partial}{\partial\theta}\ln[1-F(x_{i:m:n};\theta)]\right\}\xrightarrow{d}N(0,\gamma^2+(1-\tau)\Lambda)$$
$$=N(0,\zeta_1^2)$$

which completes the proof of Theorem 5. □

Acknowledgments The authors express their sincere thanks to Professor Udo Kamps (Editor-in-Chief) and an anonymous referee for their constructive comments and suggestions which led to a considerable improvement in the presentation of this paper. The first author thanks the National Science of Council of Taiwan (NSC Grant Number 96-2628-M-032-002-MY3) for funding this research, while the second author expresses thanks to the Natural Sciences and Engineering Research Council of Canada.

References

Balakrishnan N (2007) Progressive censoring methodology: an appraisal (with discussions). Test 16:211–296

Balakrishnan N, Aggarwala R (2000) Progressive censoring: theory, methods, and applications. Birkhäuser, Boston

Bhattacharyya GK (1985) The asymptotics of maximum likelihood and related estimators based on Type-II censored data. J Am Stat Assoc 80:398–404

Cramér H (1946) Mathematical methods of statistics. Princeton University Press, Princeton

Ferguson TS (1996) A course in large sample theory. Chapman & Hall, New York

Halperin M (1952) Maximum likelihood estimation in truncated samples. Ann Math Stat 23:226–238

Hoadley B (1971) Asymptotic properties of maximum likelihood estimators for the independent not identically distributed case. Ann Math Stat 42:1977–1991

Kong F, Fei H (1996) Limit theorems for the maximum likelihood estimate under general multiply Type II censoring. Ann Inst Stat Math 48:731–755

Louis TA (1982) Finding the observed information matrix when using the EM algorithm. J R Stat Soc Series B 44:226–233

Ng HKT, Chan PS, Balakrishnan N (2002) Estimation of parameters from progressively censored data using the EM algorithm. Comput Stat Data Anal 39:371–386

Sen PK (1976) Weak convergence of progressively censored likelihood ratio statistics and its role in asymptotic theory of life testing. Ann Stat 4:1247–1257

Serfling RJ (1980) Approximation theorems of mathematical statistics. Wiley, New York

Sethuraman J (1961) Some limit theorems for joint distributions. Sankhyā, Series A 23:379–386

Tanner MA (1993) Tools for statistical inference: observed data and data augmentation methods, 2nd edn. Springer, New York