Biostatistics (2014), **0**, 0, *pp.* 1–6
doi:10.1093/biostatistics/ODS˙Cox˙Supp

Estimating Effect of Environmental Contaminants on Women's Subfecundity for the MoBa Study Data with an Outcome-Dependent Sampling Scheme

JIELI DING

School of Mathematics and Statistics, Wuhan University, Wuhan, Hubei 430072, China

HAIBO ZHOU*

Department of Biostatistics, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599, USA
zhou@bios.unc.edu

YANYAN LIU

School of Mathematics and Statistics, Wuhan University, Wuhan, Hubei 430072, China

JIANWEN CAI

Department of Biostatistics, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599, USA

MATTHEW P. LONGNECKER

National Institute of Environmental Health Sciences, National Institute of Health, Research Triangle Park, North Carolina 27709, USA

Appendix: Proof of Asymptotic Properties

We establish the consistency of the proposed EMSELE $\hat{\xi}$ for ξ^0 by applying the lemma developed by Weaver (2001) as an extension of the result of Foutz (1977), which states the existence of a unique consistent solution to a general estimating equation. We then derive the asymptotic normal properties of EMSELE by standard methods given the consistency.

We require the following conditions:

*To whom correspondence should be addressed.

(C1) The parameter space, $\mathcal{B}$, is a compact subset of R^p. β^0 lies in the interior of $\mathcal{B}$, and the covariate space, $\mathcal{Z}$, is compact;

(C2) $\lambda_0(t)$ is strictly positive and differentiable for $t \in [0, \tau]$;

(C3) Let $S_0^{(d)}(\beta, t)$ and $s_0^{(d)}(\beta, t)$ $(d = 0, 1, 2)$ be as in Section 2.3. There exists a neighborhood D of β^0 such that $\sup_{t\in[0,\tau],\beta\in D} \|S_0^{(d)}(\beta, t) - s_0^{(d)}(\beta, t)\| \xrightarrow{P} 0, \quad d = 0, 1, 2;$

(C4) $s_0^{(d)}(\beta, t),\ d = 0, 1, 2$ are continuous functions of $\beta \in D$ uniformly in $t \in [0, \tau]$. $s_0^{(d)}(\beta, t),\ d = 0, 1, 2$ are bounded on $D \times [0, \tau]$, and $s_0^{(0)}(\beta, t)$ is bounded away from zero on $D \times [0, \tau]$. Interchanges of differentiation and integration of $s_0^{(0)}(\beta, t)$ are valid for the first and second partial derivatives with respect to β;

(C5) The matrix $A_0(\beta^0) = \int_0^\tau v_0(\beta^0, t)s_0^{(0)}(\beta^0, t)\lambda_0(t)dt$ is positive definite.

(C6) Interchanges of differentiation and integration of $f_{\beta,\Lambda_0}(t|Z)$ are valid for the first and second partial derivatives with respect to β;

(C7) The quantities $\sum_{l=1}^{K} \frac{\rho_l/\rho_0}{\pi_l^0} \nabla_{\beta_j} P_l(Z;\ \beta^0),\ j = 1, \cdots, p$ are linearly independent on $\mathcal{Z}$, i.e., if α is any p-vector such that $\sum_{j=1}^{p} \alpha_j \sum_{l=1}^{K} \frac{\rho_l/\rho_0}{\pi_l^0} \nabla_{\beta_j} P_l(Z;\ \beta^0) = 0$ for almost all $Z \in \mathcal{Z}$, then $\alpha = 0$.

Proof of Theorem 2.1: We first calculate the first and second derivatives of the profile likelihood function. We write the resulting profile likelihood function in (2.8) as:

$$
\begin{aligned}
\hat{l}(\beta, \pi) &= \hat{l}_1(\beta) + \hat{l}_2(\beta) + \hat{l}_3(\beta, \pi), \\
\hat{l}_1(\beta) &= \Delta_i \left(\beta' Z_i - \log(\sum_{l\in S_0} Y_l(T_i)e^{\beta' Z_l}) \right), \\
\hat{l}_2(\beta) &= \sum_{k=1}^{K} \sum_{i\in S_k} \log f_{\beta,\hat{\Lambda}_0}(T_i|Z_i), \\
\hat{l}_3(\beta, \pi) &= -\sum_{k=0}^{K} \sum_{i\in S_k} \log \left(n_0(1 + \sum_{l=1}^{K} \frac{n_l}{n_0 \pi_l} P_l(Z_i; \beta)) \right) - \sum_{k=1}^{K} n_k \log \pi_k.
\end{aligned} \tag{A.1}
$$

The first derivatives of the profile likelihood function are:

$$\frac{\partial \hat{l}_1(\beta)}{\partial \beta} = \sum_{i \in S_0} \int_0^\tau \left(Z_i - \frac{S^{(1)}(\beta,t)}{S^{(0)}(\beta,t)} \right) dN_i(t),$$

$$\frac{\partial \hat{l}_2(\beta)}{\partial \beta} = \sum_{k=1}^{K} \sum_{i \in S_k} \frac{\nabla_\beta f_{\beta,\hat{\Lambda}_0}(T_i|Z_i)}{f_{\beta,\hat{\Lambda}_0}(T_i|Z_i)},$$

$$\frac{\partial \hat{l}_3(\beta,\pi)}{\partial \beta} = -\sum_{k=0}^{K} \sum_{i \in S_k} \frac{\sum_{l=1}^K \frac{n_l}{n_0 \pi_l} \nabla_\beta P_l(Z_i;\beta)}{1 + \sum_{l=1}^K \frac{n_l}{n_0 \pi_l} P_l(Z_i;\beta)},$$

$$\frac{\partial \hat{l}_3(\beta,\pi)}{\partial \pi_m} = \sum_{k=0}^{K} \sum_{i \in S_k} \frac{\frac{n_m}{n_0 \pi_m^2} P_m(Z_i;\beta)}{1 + \sum_{l=1}^K \frac{n_l}{n_0 \pi_l} P_l(Z_i;\beta)} - \frac{n_m}{\pi_m}, \qquad m = 1, \cdots, K. \tag{A.2}$$

The second derivatives of the profile likelihood function are:

$$-\frac{\partial^2 \hat{l}_1(\beta)}{\partial \beta \partial \beta'} = \sum_{i \in S_0} \int_0^\tau \left(\frac{S^{(2)}(\beta,t)}{S^{(0)}(\beta,t)} - \left(\frac{S^{(1)}(\beta,t)}{S^{(0)}(\beta,t)} \right)^{\otimes 2} \right) dN_i(t),$$

$$-\frac{\partial^2 \hat{l}_2(\beta)}{\partial \beta \partial \beta'} = -\sum_{k=1}^{K} \sum_{i \in S_k} \left(\frac{\nabla_\beta^2 f_{\beta,\hat{\Lambda}_0}(T_i|Z_i)}{f_{\beta,\hat{\Lambda}_0}(T_i|Z_i)} - \left(\frac{\nabla_\beta f_{\beta,\hat{\Lambda}_0}(T_i|Z_i)}{f_{\beta,\hat{\Lambda}_0}(T_i|Z_i)} \right)^{\otimes 2} \right),$$

$$-\frac{\partial^2 \hat{l}_3(\beta,\pi)}{\partial \beta \partial \beta'} = \sum_{k=0}^{K} \sum_{i \in S_k} \left(\frac{\sum_{l=1}^K \frac{n_l}{n_0 \pi_l} \nabla_\beta^2 P_l(Z_i;\beta)}{1 + \sum_{l=1}^K \frac{n_l}{n_0 \tilde{\pi}_l} P_l(Z_i;\beta)} - \frac{\left(\sum_{l=1}^K \frac{n_l}{n_0 \pi_l} \nabla_\beta P_l(Z_i;\beta) \right)^{\otimes 2}}{\left(1 + \sum_{l=1}^K \frac{n_l}{n_0 \pi_l} P_l(Z_i;\beta) \right)^2} \right),$$

and

$$-\frac{\partial^2 \hat{l}_3(\beta,\pi)}{\partial \beta \partial \pi_m'} = \sum_{k=0}^{K} \sum_{i \in S_k} \frac{-\frac{n_m}{n_0 \pi_m^2} \nabla_\beta P_m(Z_i;\beta)}{1 + \sum_{l=1}^K \frac{n_l}{n_0 \pi_l} P_l(Z_i;\beta)} + \sum_{k=0}^{K} \sum_{i \in S_k} \frac{\frac{n_m}{n_0 \pi_m^2} P_m(Z_i;\beta) \left(\sum_{l=1}^K \frac{n_l}{n_0 \pi_l} \nabla_\beta P_l(Z_i;\beta) \right)}{\left(1 + \sum_{l=1}^K \frac{n_l}{n_0 \pi_l} P_l(Z_i;\beta) \right)^2},$$

$$-\frac{\partial \hat{l}_3(\beta,\pi)}{\partial \pi_m \partial \pi_r} = -\sum_{k=0}^{K} \sum_{i \in S_k} \frac{\frac{n_m}{n_0 \pi_m^2} P_m(Z_i;\beta) \frac{n_r}{n_0 \pi_r^2} P_r(Z_i;\beta)}{\left(1 + \sum_{l=1}^K \frac{n_l}{n_0 \pi_l} P_l(Z_i;\beta) \right)^2} + \sum_{k=0}^{K} \sum_{i \in S_k} \frac{\frac{2 n_m}{n_0 \pi_m^3} P_m(Z_i;\beta) \delta_{mr}}{1 + \sum_{l=1}^K \frac{n_l}{n_0 \pi_l} P_l(Z_i;\beta)} - \frac{n_m}{\pi_m^2} \delta_{mr}, \tag{A.3}$$

where $\delta_{mr} = 1$ if $m = r$ and 0 otherwise, for $m, r = 1, \cdots, K$.

By a straightforward extension of the classical results on Cox's model (Andersen and Gill, 1982; Kalbfleisch and Prentice, 2002), we can obtain that $\frac{1}{n} \frac{\partial \hat{l}_1(\beta)}{\partial \beta} \xrightarrow{P} 0$, $-\frac{1}{n} \frac{\partial^2 \hat{l}_1(\beta)}{\partial \beta \partial \beta'} \xrightarrow{P} \rho_0 A_0(\beta)$ uniformly for β in a neighborhood of β^0, and $\frac{1}{\sqrt{n}} \frac{\partial \hat{l}_1(\beta^0)}{\partial \beta} \xrightarrow{d} N(0, \rho_0 A_0(\beta^0))$, where $\rho_0 = \lim_{n \to \infty} n_0/n$, and $A_0(\beta)$ is described in Section 2.3. Thus, according to the law of large numbers, we can show that

$$
\begin{aligned}
\frac{1}{n}\frac{\partial \hat{l}(\xi)}{\partial \beta} &\xrightarrow{P} \sum_{k=1}^{K}\rho_k E_k\left(\frac{\nabla_\beta f_{\beta,\hat{\Lambda}_0}(T|Z)}{f_{\beta,\hat{\Lambda}_0}(T|Z)}\right) - \sum_{k=0}^{K}\rho_k E_k\left(\frac{\sum_{l=1}^{K}\frac{\rho_l}{\rho_0\pi_l}\nabla_\beta P_l(Z;\beta)}{1+\sum_{l=1}^{K}\frac{\rho_l}{\rho_0\pi_l}P_l(Z;\beta)}\right) \equiv s_1(\xi),\\
\frac{1}{n}\frac{\partial \hat{l}(\xi)}{\partial \pi} &\xrightarrow{P} \sum_{k=0}^{K}\rho_k E_k\begin{pmatrix}\frac{\frac{\rho_1}{\rho_0\pi_1^2}P_1(Z;\beta)}{1+\sum_{l=1}^{K}\frac{\rho_l}{\rho_0\pi_l}P_l(Z;\beta)}\\ \vdots \\ \frac{\frac{\rho_K}{\rho_0\pi_K^2}P_K(Z;\beta)}{1+\sum_{l=1}^{K}\frac{\rho_l}{\rho_0\pi_l}P_l(Z;\beta)}\end{pmatrix} - \begin{pmatrix}\frac{\rho_1}{\pi_1}\\ \vdots \\ \frac{\rho_K}{\pi_K}\end{pmatrix} \equiv s_2(\xi),
\end{aligned}
\tag{A.4}
$$

where $\rho_0 = \lim_{n\to\infty} n_0/n,\ \rho_k = \lim_{n\to\infty} n_k/n$. Then we have $\frac{1}{n}\frac{\partial \hat{l}(\xi)}{\partial \xi}$ converges to $s(\xi) \equiv \begin{pmatrix} s_1(\xi) \\ s_2(\xi)\end{pmatrix}$ in probability. When evaluate at the true value ξ^0, we can show $s(\xi^0) \xrightarrow{P} 0$.

Furthermore, by the uniform convergence theorem established by Jennich (1969), we can obtain that $-\frac{1}{n}\frac{\partial^2 \hat{l}(\xi)}{\partial\xi\partial\xi'}$ converges to $J(\xi)$ in probability uniformly for ξ in a neighborhood around ξ^0, where $J(\xi) = \begin{pmatrix} J_{11}(\xi) & J_{12}(\xi) \\ (J_{12}(\xi))' & J_{22}(\xi)\end{pmatrix}$ with

$$
\begin{aligned}
J_{11}(\xi) &= \rho_0 A_0(\beta) + \sum_{k=1}^{K}\rho_k E_k\left(-\nabla_\beta^2 \log f_{\beta,\hat{\Lambda}_0}(T|Z)\right) + \sum_{k=0}^{K}\rho_k E_k\left(\nabla_\beta^2 \log(1+\sum_{l=1}^{K}\frac{\rho_l}{\rho_0\pi_l}P_l(Z;\beta))\right),\\
J_{22}(\xi) &= \sum_{k=0}^{K}\rho_k E_k\left(\nabla_\pi^2 \log(1+\sum_{l=1}^{K}\frac{\rho_l}{\rho_0\pi_l}P_l(Z;\beta))\right) + \begin{pmatrix}\rho_1/\pi_1^2 & & \\ & \ddots & \\ & & \rho_K/\pi_K^2\end{pmatrix},\\
J_{12}(\xi) &= \sum_{k=0}^{K}\rho_k E_k\left(\nabla_{\beta\pi}^2 \log(1+\sum_{l=1}^{K}\frac{\rho_l}{\rho_0\pi_l}P_l(Z;\beta))\right).
\end{aligned}
$$

By Assumption (C7), it also can be shown that $J(\xi^0)$ is invertible. Then, the convergence of $\hat{\xi}$ is achieved by applying Lemma 3.3 in Weaver (2001).

Given the consistency result we obtain, we can derive the asymptotic normal properties of EMSELE by standard methods. By using the first-order Taylor series expansion of $\frac{\partial l_P(\xi)}{\partial \xi}$ at ξ^0, we have

$$
\sqrt{n}(\hat{\xi}-\xi^0) = \left(-\frac{1}{n}\frac{\partial^2 \hat{l}(\xi^*)}{\partial\xi\partial\xi'}\right)^{-1}\left(\begin{pmatrix}\frac{1}{\sqrt{n}}\frac{\partial \hat{l}_1(\beta^0)}{\partial\beta}\\ 0\end{pmatrix} + \begin{pmatrix}\frac{1}{\sqrt{n}}\frac{\partial \hat{l}_2(\beta^0)}{\partial\beta}\\ 0\end{pmatrix} + \frac{1}{\sqrt{n}}\frac{\partial \hat{l}_3(\xi^0)}{\partial\xi}\right), \tag{A.5}
$$

where ξ^* lies on the line between $\hat{\xi}$ and ξ^0. Using the consistency of $\hat{\xi}$, it is obvious to conclude that $\left(-\frac{1}{n}\frac{\partial^2 l_P(\xi^*)}{\partial\xi\partial\xi'}\right)^{-1} \xrightarrow{P} J^{-1}(\xi^0)$ as $n \to \infty$. It follows the Central Limit Theorem that

$$\begin{pmatrix} \frac{1}{\sqrt{n}}\frac{\partial \hat{l}_1(\beta^0)}{\partial\beta} \\ 0 \end{pmatrix} \xrightarrow{d} N\left(0,\ \begin{pmatrix}\rho_0 A_0(\beta) & 0 \\ 0 & 0\end{pmatrix}\right),$$

$$\begin{pmatrix} \frac{1}{\sqrt{n}}\frac{\partial \hat{l}_2(\beta^0)}{\partial\beta} \\ 0 \end{pmatrix} \xrightarrow{d} N\left(0,\ \begin{pmatrix}\sum_{k=1}^{K}\rho_k Var_k\left(e(\beta)\right) & 0 \\ 0 & 0\end{pmatrix}\right), \tag{A.6}$$

and

$$\frac{1}{\sqrt{n}}\frac{\partial \hat{l}_3(\xi^0)}{\partial\xi} \xrightarrow{d} N\left(0,\ \Sigma_3(\xi^0)\right), \tag{A.7}$$

with

$$\Sigma_3(\xi) = \sum_{k=0}^{K}\rho_k Var_k \begin{pmatrix} -\frac{\sum_{l=1}^{K}\frac{\rho_l/\rho_0}{\pi_l}\nabla_\beta P_l(Z;\beta)}{1+\sum_{l=1}^{K}\frac{\rho_l/\rho_0}{\pi_l}P_l(Z;\beta)} \\ \frac{\frac{\rho_1/\rho_0}{\pi_1^2}P_1(Z;\beta)}{1+\sum_{l=1}^{K}\frac{\rho_l/\rho_0}{\pi_l}P_l(Z;\beta)} - \frac{\rho_1}{\pi_1} \\ \vdots \\ \frac{\frac{\rho_K/\rho_0}{\pi_K^2}P_K(Z;\beta)}{1+\sum_{l=1}^{K}\frac{\rho_l/\rho_0}{\pi_l}P_l(Z;\beta)} - \frac{\rho_K}{\pi_K} \end{pmatrix}.$$

The asymptotic property of $\hat{\xi}$ holds by Slutsky's theorem. Finally, by consistency results and the continuous mapping theorem, a consistent estimator for the asymptotic covariance matrix $\Sigma(\xi^0)$ is $\hat{J}^{-1}(\hat{\xi})(\hat{\Sigma}_1(\hat{\xi}) + \hat{\Sigma}_2(\hat{\xi}) + \hat{\Sigma}_3(\hat{\xi}))\hat{J}^{-1}(\hat{\xi})$, where $\hat{J}$, $\hat{\Sigma}_1$ $\hat{\Sigma}_2$ and $\hat{\Sigma}_3$ are obtained by replacing the large-sample quantities in J, Σ_1 Σ_2 and Σ_3 with their corresponding small-sample quantities.

References

Andersen, P., and Gill, R. (1982). Cox's Regression Model for Counting Processes: A Large Sample Study. *The Annals of Statistics* **10**, 1100–1120.

Foutz, R. V. (1977). On the Unique Consistent Solution to the Likelihood Equations. *Journal of the American Statistical Association* **72**, 147–148.

JENNRICH, R. T. (1969). Asymptotic Properties of Non-Linear Least Squares Estimators. *The Annals of Mathematical Statistics* **40**, 633–643.

KALBFLEISCH, J. D., AND PRENCTICE, R. L. *The Statistical Analysis of Failure Time Data*. John Wiley & Sons, Inc, 2002.

WEAVER, M. A. Semiparametric Methods for Continuous Outcome Regression Models with Covariate Data from an Outcome Dependent Subsample. Ph.D. Thesis, University Of North Carolina At Chapel Hill, 2001.