Fang and Wu *Journal of Inequalities and Applications* (2015) 2015:379
DOI 10.1186/s13660-015-0903-1

Journal of Inequalities and Applications
a SpringerOpen Journal

RESEARCH Open Access

Empirical likelihood inference of parameters in nonlinear EV models with censored data

Liandi Fang and Yongfeng Wu*

*Correspondence: wyfwyf@126.com
College of Mathematics and Computer Science, Tongling University, Tongling, 244000, China

Abstract

In this paper, the authors consider nonlinear error-in-variables regression models with right censored data. Based on the validation data, the authors use the Kaplan-Meier estimate and kernel estimate to repair the censored response variable Y and explanatory X with measurement error. Based on the repaired data, the authors introduce an auxiliary vector suitable to define an estimated empirical log-likelihood function of the unknown parameter which has an asymptotic weighted sum of the χ_1^2 variables. Using the result the authors can construct the asymptotic confidence regions of β. But the method of estimating weights will reduce the precision of the confidence regions. Further, the authors adjust the preceding log-likelihood, and it is shown that the adjusted empirical log-likelihood has the asymptotic standard χ_1^2 distribution. The result can be used to construct the confidence regions of β.

MSC: 62G05; 62J02

Keywords: empirical log-likelihood; censored data; error-in-variables; validation data

1 Introduction

Consider the following nonlinear model:

$$Y = g(X, \beta) + \varepsilon, \qquad \widetilde{X} = \phi(X, e), \tag{1.1}$$

where X is p-variate explanatory variable, Y is a scalar response variable, and where $\beta = (\beta_1, \beta_2, \ldots, \beta_p)^\tau$ is a $p \times 1$ column vector of the unknown regression parameter, $g(\cdot)$ is a known measurable function, and ε is a random statistical error. In this model, X is an explanatory variable which cannot be observed directly, and $\widetilde{X}$ is the observable substitute variable of X, where e is a measurement error, $\phi(\cdot)$ is an unconditional known function. Usually, there is a complicated relationship between $\widetilde{X}$ and X. However, this situation presents serious difficulties toward obtaining a correct statistical analysis. One solution is to use validation data, and the solution has gained much attention by many scholars in recent years. For example: Sepanki and Lee [1] considered error-in-covariable nonlinear models with the help of validation data. Wang [2–4] considered the empirical likelihood inference of the error-in-covariable linear model and partially linear model based on validation data. Xue [5] used the validation data to explore the empirical likelihood inference of nonlinear semiparasitic error-in-variable models. Fang and Hu [6] considered the error-in-response to be nonlinear addressing the dimension reduction estimation of β with validation data.

In practice, the response Y's may be censored by the censoring variable C. So we cannot observe (X_i, Y_i), but we may observe

$$(\widetilde{X}_i, Z_i, \delta_i), \quad i = 1, \ldots, n,$$

with $Z_i = \min(Y_i, C_i)$, $\delta_i = I(Y_i \le C_i)$, where $I(\cdot)$ is the indicator function and $\delta = 0$ if Y_i is censored, otherwise $\delta = 1$. $(\widetilde{X}_i, Z_i, \delta_i)$, $i = 1, \ldots, n$, are i.i.d. random samples from $(\widetilde{X}, Z, \delta)$. Suppose that $\widetilde{X}$, Z, are given and δ is independent of Y. In fact, regression models with censored data have been researched in much of the literature. For example, Buckley and James [7] made an unbiased change for data and established a new regression model. Based on [7], Li and Wang [8] supposed the censored variable to be independent of the covariable and applied a new data change to proceed with the empirical likelihood inference in linear models. Cheng *et al.* [9] developed the empirical likelihood inference in nonlinear models with a right censored response based on the validation data.

In this paper, we consider the model (1.1), where the response Y is randomly censored, and explanatory X with measurement error. We try to construct the confidence region for the parameter β. Similarly, we propose to use the empirical likelihood method to construct the confidence region for β. A typical nonparametric approach to this problem generally includes the following steps: (1) repair the incomplete data with the help of validation data and derive an $\hat{\beta}$ to estimate β, (2) define the empirical likelihood function, (3) invert the confidence region by the limiting χ_1^2 distribution. To complete these steps, we will carry out the following approach. First, based on the validation data, we use the Kaplan-Meier estimate and the kernel estimate to repair the censored response variable Y and explanatory X, which has been measured erroneously. Second, based on the repaired data, we introduce an auxiliary vector suitable to define an estimated empirical log-likelihood function of the unknown parameter which has an asymptotic weighted sum of the χ_1^2 variables. Because the weights are unknown, we can give the corresponding estimators of the weights. Finally, using the result we can construct the asymptotic confidence regions of β. But the method of estimating weights will reduce the precision of the confidence regions. Further, we adjust the preceding log-likelihood, and it is shown that the adjusted empirical log-likelihood has the asymptotic standard χ_1^2 distribution.

2 Results and discussion

2.1 Definition of estimated empirical function

Suppose the primary data set $(\widetilde{X}_i, Z_i, \delta_i)_{i=1}^n$ is independent of the validation data set $(\widetilde{X}_j, X_j)_{j=n+1}^{n+m}$, $E[\varepsilon|X] = 0$, and $E[\varepsilon|\widetilde{X}] = 0$. Further we suppose $F(\cdot)$ and $G(\cdot)$ are the density functions of Y and C, respectively. Because $\{Y_i\}$ is censored, and the completely observed variable Z_i has different expectations to Y_i, we cannot directly use the general method to estimate β. When G is known, we define

$$Y_{iG} = \frac{\delta_i Z_i}{1 - G(Z_i-)}.$$

Denote $m(\widetilde{X},\beta) = E[g(X,\beta)|\widetilde{X}]$. It is well known that $E[Y_{iG}|\widetilde{X}_i] = E[Y_i|\widetilde{X}_i] = m(\widetilde{X}_i,\beta)$. Then based on the complete observed data, from model (1.1) can be switched to the following models:

$$Y_{iG} = m(\widetilde{X}_i,\beta) + \eta_i, \tag{2.1}$$

where $\eta_i = Y_{iG} - m(\widetilde{X}_i,\beta)$. Write

$$g^{(1)}(X,\beta) = \left(\frac{\partial}{\partial\beta}g(X,\beta)\right)^\tau = \left(\left(\frac{\partial}{\partial\beta_1},\ldots,\frac{\partial}{\partial\beta_p}\right)g(X,\beta)\right)^\tau,$$
$$m^{(1)}(\widetilde{X}_i,\beta) = \frac{\partial}{\partial\beta}m(\widetilde{X}_i,\beta) = E\big[g^{(1)}(X,\beta)|\widetilde{X}_i\big].$$

Suppose G is known, introduce an auxiliary random vector

$$W_i(\beta) = m^{(1)}(\widetilde{X}_i,\beta)\big[Y_{iG} - m(\widetilde{X}_i,\beta)\big]. \tag{2.2}$$

It is easily shown that $E[W_i(\beta)] = E\{m^{(1)}(\widetilde{X}_i,\beta)E[\eta_i|\widetilde{X}_i]\} = 0$, if β is the true value of the parameter.

In practice, $m(\widetilde{X},\beta)$, $m^{(1)}(\widetilde{X},\beta)$ and G are usually unknown. For establishing the empirical likelihood, their estimators need to be given first. To do this, for G, we employ the Kaplan-Meier estimator

$$\widehat{G}(t) = 1 - \prod_{i=1}^{n}\left[\frac{n-i}{n-i+1}\right]^{I[Z_{(i)}\le t,\delta_{(i)}=0]},$$

where $Z_{(1)} \le Z_{(2)} \le \cdots \le Z_{(n)}$ are the order statistics of $Z_1, Z_2, \ldots, Z_n$, and $\delta_{(i)}$ are the indicators associated with $Z_{(i)}$, $i = 1,\ldots,n$. Suppose

$$\widehat{R}_m(\tilde{x},\beta) = \frac{1}{mh^p}\sum_{j=n+1}^{n+m} g(X_j,\beta)K\left(\frac{\widetilde{X}_j - \tilde{x}}{h}\right),$$
$$\hat{f}_m(\tilde{x}) = \frac{1}{mh^p}\sum_{j=n+1}^{n+m} K\left(\frac{\widetilde{X}_j - \tilde{x}}{h}\right), \qquad \widehat{R}_m^{(1)}(\tilde{x},\beta) = \frac{\partial}{\partial\beta}\widehat{R}_m(\tilde{x},\beta).$$

Here $K(\cdot)$ is a kernel function and $h = h_m$ is a bandwidth tending to 0. Denote $\hat{f}_{b_m}(\tilde{x}) = \max\{\hat{f}_m(\tilde{x}), b_m\}$, where b_m is bounded zero positive numbers. Using the validation data, we define the blocked estimators of $m(\widetilde{X},\beta)$ and $m^{(1)}(\widetilde{X},\beta)$ as follows:

$$\hat{m}(\tilde{x},\beta) = \frac{\widehat{R}_m(\tilde{x},\beta)}{\hat{f}_{b_m}(\tilde{x})}, \qquad \hat{m}^{(1)}(\tilde{x},\beta) = \frac{\widehat{R}_m^{(1)}(\tilde{x},\beta)}{\hat{f}_{b_m}(\tilde{x})}.$$

Use their estimators $\hat{m}(\widetilde{X},\beta)$, $\hat{m}^{(1)}(\widetilde{X},\beta)$, and $\widehat{G}(\cdot)$ to replace the unknown functions $m(\widetilde{X},\beta)$, $m^{(1)}(\widetilde{X},\beta)$, and G in (2.2), and write

$$\widehat{W}_i(\beta) = \hat{m}^{(1)}(\widetilde{X}_i,\beta)\big[Y_{i\widehat{G}} - \hat{m}(\widetilde{X}_i,\beta)\big]. \tag{2.3}$$

It is easily proved that $E[\widehat{W}_i(\beta)] = o(1)$ when β is the true value. Using this, an estimated empirical log-likelihood-ratio function is defined as

$$\hat{l}(\beta) = -2\max\left\{\sum_{i=1}^{n}\log(np_i)\middle| p_i \geq 0, \sum_{i=1}^{n} p_i \widehat{W}_i(\beta) = 0, \sum_{i=1}^{n} p_i = 1\right\}.$$

By introducing the Lagrange multipliers, the most fines value of p_i is $p_i = n^{-1}(1 + \lambda^\tau \widehat{W})^{-1}$, where λ is determined by

$$\frac{1}{n}\sum_{i=1}^{n}\frac{\widehat{W}_i(\beta)}{1 + \lambda^\tau \widehat{W}_i(\beta)} = 0. \tag{2.4}$$

So $\hat{l}(\beta)$ can be represented as

$$\hat{l}(\beta) = 2\sum_{i=1}^{n}\log\bigl(1 + \lambda^\tau \widehat{W}_i(\beta)\bigr). \tag{2.5}$$

In succession, we define the β's estimator by minimizing

$$\hat{\beta} = \arg\min_{\beta}\frac{1}{2}\sum_{i=1}^{n}\bigl[Y_{i\widehat{G}} - \hat{m}(\widetilde{X}_i, \beta)\bigr]^2. \tag{2.6}$$

2.2 Construction of confidence region

Throughout this section, we use $c > o$ to represent any constant which does not rely on n and m and may take different values for each appearance. Let M^k be a class of all continuous function classes in R^p $(k > p)$ or subdomains of R^p which make the partial derivatives $\frac{\partial^{i_1}}{\partial x_1^{i_1}} \cdot \frac{\partial^{i_2}}{\partial x_2^{i_2}} \cdots \frac{\partial^{i_p}}{\partial x_p^{i_p}}\varphi(x_1, \ldots, x_p)$ uniformly bounded for $o < i_i + \cdots + i_p \leq k$.

To obtain our result, we need to list the following conditions.

Condition C

- C_1. $E\|X\|^2 < \infty$, $EY^2 < \infty$.
- C_2. $g(x, \beta)$ has bounded continuous partial derivatives up to order two in Γ where Γ is the bounded support of x.
- C_3. For some k $(> p)$, there is $m(\tilde{x}, \beta) \in M^k$, $m_s^{(1)}(\tilde{x}, \beta) \in M^k$, $s = 1, 2, \ldots, p$.
- C_4. $K(u)$ is a bounded nonnegative kernel function of order k $(k > p)$ with bounded support.
- C_5. The density of $\widetilde{X}$, say $f(\tilde{x})$, satisfies
  - (i) $f(\tilde{x}) \in M^k$,
  - (ii) there exists a positive constant sequence b_m, such that $\sqrt{m}P\{f(\tilde{x}) < b_m\} \to 0$ as $m \to \infty$.
- C_6. $mh^{2p}b_m^4 \to \infty$, $mh^{2k}b_m^{-2} \to 0$, $h^{k-\frac{1}{2}p}b_m \to 0$.
- C_7. $\sup_{(\tilde{x},x)} E[e^2|\widetilde{X} = \tilde{x}, X = x] < \infty$.
- C_8. $\frac{n}{m} \to \gamma$, where $\gamma > 0$ is a nonnegative constant.
- C_9. For any $0 \leq s < \infty$, there exists $\Gamma_1(s) = E[m^{(1)}(\widetilde{X}, \beta)I[s < Y]]$.
- C_{10}. (i) For any $s \leq \tau_Q = \inf t : Q(t) = 1$, $G(s)$, and $F(s)$ has no common jumps where $Q(t) = P(Z \leq t)$,

(ii) $E\frac{\|g^{(1)}(X,\beta)\|Y}{[(1-G(Y))(1-F(Y))]^{\frac{1}{2}}} < \infty$,

(iii) $\int_0^{\tau_Q} \|H(s)\| \frac{[1-F(s)]}{[1-F(s-)][1-G(s)]}\, dG(s) < \infty$, where $H(s) = \frac{E[m^{(1)}(\widetilde{X},\beta)Y_G I[s<Z]]}{[1-F(s-)][1-G(s)]}$.

C_{11}. $\Sigma_0(\beta) = E[Y_G - m(\widetilde{X},\beta)]^2 m^{(1)}(\widetilde{X},\beta)(m^{(1)}(\widetilde{X},\beta))^\tau = E[W_1(\beta)(W_1(\beta))^\tau]$, $\Xi = Em^{(1)}(\widetilde{X},\beta)(m^{(1)}(\widetilde{X},\beta))^\tau$, $\Sigma_1(\beta) = \int_0^{+\infty} H(s)H^T(s)\bar{F}(s-)(1-\Lambda^G(s))\, dG(S)$, where $\Lambda^G(s) = \int_{-\infty}^t \bar{G}^{-1}(s-)\, dG(s)$, and $\Sigma_0(\beta)$, $\Sigma_1(\beta)$, and Ξ are all positive definite matrices.

Remark 2.1 These conditions are some usual assumptions for studying the semiparametric model and can be satisfied. Here, condition C_6 is only explained, h was taken to be $h = c_2 m^{-\frac{1}{p+k}}$ if $b_m = c_1 m^{\frac{p-k}{4(p+k)}} \log m$, where c_1 and c_2 are positive constants.

Theorem 2.1 *Under Condition* C, *if* β *is the true value of the parameter, we have*

$$\sqrt{n}(\hat{\beta} - \beta) \xrightarrow{L} N\bigl(0, \Xi^{-1}\Sigma\Xi^{-1}\bigr),$$

where Ξ *is defined in condition* C_{11}, $\Sigma = \Sigma(\beta) = \Sigma_0(\beta) - \Sigma_1(\beta) + \Sigma_2(\beta)$, *with* $\Sigma_0(\beta)$, $\Sigma_1(\beta)$ *being defined in condition* C_{11}, *and* $\Sigma_2(\beta) = \gamma E\{m^{(1)}(\widetilde{X},\beta)(m^{(1)}(\widetilde{X},\beta))^\tau[m(\widetilde{X},\beta) - g(X,\beta)]^2\}$ *where* γ *is defined in condition* C_8.

Theorem 2.2 *Under Condition* C, *if* β *is the true value of the parameter, we have*

$$\hat{l}(\beta) \xrightarrow{L} \omega_1\chi^2_{1,1} + \omega_2\chi^2_{1,2} + \cdots + \omega_p\chi^2_{1,p}, \tag{2.7}$$

where $\chi^2_{1,i}$ $(1 \le i \le p)$ *are independent standard* χ^2 *random variables with* 1 *degree of freedom, and* ω_i $(1 \le i \le p)$ *are the eigenvalues of* $D(\beta) = \Sigma_0^{-1}(\beta)\Sigma(\beta)$.

In order to use the result of Theorem 2.2 to construct the confidence regions of β, the corresponding estimators of the unknown weights ω_i $(1 \le i \le p)$ must be given. Denote by $\widehat{F}$ the *Kaplan-Meier* estimator of F, and let $Q_n(s) = \frac{1}{n}\sum_{i=1}^n I(Z_i \le s)$. Using the plug-in method, we give the following notations:

$\widehat{H}_n(s) = \frac{\frac{1}{n}\sum_{i=1}^n \hat{m}^{(1)}(\widetilde{X}_i,\hat{\beta})Y_{i\widehat{G}} I(s<Z_i)}{(1-\widehat{G}(s))(1-\widehat{F}(s-))}$,

$\Lambda_n^{\widehat{G}}(s) = \int_{-\infty}^s \frac{1}{1-\widehat{G}(t-)}\, d\widehat{G}(t) = \frac{1}{n}\sum_{i=1}^n \frac{(1-\delta_i)I(Z_i\le t)}{(1-\widehat{G}(t-))(1-Q_n(t-))}$,

$\widehat{\Sigma}_0(\hat{\beta}) = \frac{1}{n}\sum_{i=1}^n [Y_{i\widehat{G}} - \hat{m}(\widetilde{X}_i,\hat{\beta})]^2 \hat{m}^{(1)}(\widetilde{X}_i,\hat{\beta})(\hat{m}^{(1)}(\widetilde{X}_i,\hat{\beta}))^\tau = \frac{1}{n}\sum_{i=1}^n \widehat{W}_i(\beta)\widehat{W}_i^T(\beta)$,

$\widehat{\Sigma}_1(\hat{\beta}) = \frac{1}{n}\sum_{i=1}^n (1-\delta_i)\widehat{H}_n(Z_i)\widehat{H}_n^\tau(Z_i)(1-\Lambda_n^{\widehat{G}}(Z_i))$,

$\widehat{\Sigma}_2(\hat{\beta}) = \frac{\gamma}{m}\sum_{j=n+1}^{n+m} \hat{m}^{(1)}(\widetilde{X}_j,\hat{\beta})(\hat{m}^{(1)}(\widetilde{X}_j,\hat{\beta}))^\tau[\hat{m}(\widetilde{X}_j,\hat{\beta}) - g(X_j,\beta)]^2$,

$\widehat{\Sigma}(\hat{\beta}) = \widehat{\Sigma}_0(\hat{\beta}) - \widehat{\Sigma}_1(\beta) + \widehat{\Sigma}_2(\hat{\beta})$.

From this, we can infer that $\hat{\omega}_i$ $(i = 1,2,\ldots,p)$ (the eigenvalues of $\widehat{D}(\hat{\beta}) = \widehat{\Sigma}_0^{-1}(\hat{\beta})\widehat{\Sigma}(\hat{\beta})$) are the corresponding estimators of ω_i $(i = 1,2,\ldots,p)$. Denote $\hat{s} = \hat{\omega}_1\chi^2_{1,1} + \hat{\omega}_2\chi^2_{1,2} + \cdots + \hat{\omega}_p\chi^2_{1,p}$. $B_f(\cdot)$ expresses the conditional distribution of $\hat{s}$ given $(\widetilde{X}_i, Z_i, \delta_i)_{i=1}^n$ and $(\widetilde{X}_j, X_j)_{j=n+1}^{n+m}$. Let $\hat{c}_\alpha$ be the $1-\alpha$ fractile of $B_f(\cdot)$, and the $1-\alpha$ confidence region for β is

$$\hat{I}_\alpha(\tilde{\beta}) = \bigl\{\tilde{\beta} : \hat{l}(\tilde{\beta}) \le \hat{c}_\alpha\bigr\}.$$

Actually, it is convenient to obtain the conditional distribution $B_f(\cdot)$, we can generate independent samples $\chi^2_{1,1},\ldots,\chi^2_{1,p}$ from the χ^2_1 distribution and then get it through the Monte Carlo simulation method.

2.3 Definition of adjusted empirical function

Applying the result of Theorem 2.2 to construct the confidence region of β, we need to estimate the weights ω_i, which will reduce the accuracy of the confidence region. So we need to adjust the empirical likelihood function. Let $r(\beta) = \frac{p}{\operatorname{tr}\{D(\beta)\}}$, according to the result given by Rao and Scott [10], $r(\beta)\sum_{i=1}^{p}\omega_i\chi_{1,i}^2$ has an asymptotic standard χ^2-distribution with p degrees of freedom. From Theorem 2.1 and the consistency of $\widehat{\Sigma}(\hat{\beta})$ and $\widehat{\Sigma}_0(\hat{\beta})$, we can find that $\hat{r}(\hat{\beta})\hat{l}(\beta)$ also has an asymptotic standard χ^2-distribution with p degrees of freedom, where $\hat{r}(\hat{\beta}) = \frac{p}{\operatorname{tr}\{\widehat{D}(\hat{\beta})\}} = \frac{p}{\operatorname{tr}\{\widehat{\Sigma}_0^{-1}(\hat{\beta})\widehat{\Sigma}(\hat{\beta})\}}$. For improving the approximation accuracy, we replace $\hat{\beta}$ with β in $\hat{r}(\hat{\beta})$ and the accuracy will depend on the value of ω_i. We can refine the result given by Rao and Scott [10], and then we give an adjusted empirical log-likelihood ratio. Denote $\widehat{A}(\beta) = \{\sum_{i=1}^{n}\widehat{W}_i(\beta)\}\{\sum_{i=1}^{n}\widehat{W}_i(\beta)\}^{\tau}$. If we replace $\widehat{\Sigma}(\beta)$ with $\widehat{A}(\beta)$ in $\hat{r}(\beta)$, we will get a new adjustment factor $\hat{\rho}(\beta) = \frac{\operatorname{tr}\{\widehat{\Sigma}^{-1}(\beta)\widehat{A}(\beta)\}}{\operatorname{tr}\{\widehat{\Sigma}_0^{-1}(\beta)\widehat{A}(\beta)\}}$, then the adjusted empirical log-likelihood ratio function can be defined as

$$\hat{l}_{\mathrm{ad}}(\beta) = \hat{\rho}(\beta)\hat{l}(\beta).$$

Theorem 2.3 *Under Condition* C, *if* β *is the true value of the parameter, we have*

$$\hat{l}_{\mathrm{ad}}(\beta) \xrightarrow{L} \chi_p^2.$$

Based on Theorem 2.3, $\hat{l}_{\mathrm{ad}}(\beta)$ can be used to construct a confidence region for β,

$$\hat{I}_{\mathrm{ad}}(\tilde{\beta}) = \{\tilde{\beta} : \hat{l}_{\mathrm{ad}}(\beta) \le c_\alpha\},$$

where $P(\chi_p^2 \le c_\alpha) = 1 - \alpha$, and $P\{\beta \in \hat{I}_{\mathrm{ad}}(\tilde{\beta})\} = 1 - \alpha + o(1)$.

3 Proofs of theorems

Before the proofs of the theorems, we introduce some preliminary results.

Lemma 3.1 *Under Condition* C, *for any* $1 \le i \le n$, *we have*
(i) $E[(\hat{m}(\widetilde{X}_i,\beta) - m(\widetilde{X}_i,\beta))^2|\widetilde{X}_i] \le c(mh^p b_m^2)^{-1} + ch^{2k}b_m^{-2} + cI[F(\widetilde{X}_i) < 2b_m]$;
(ii) $E[(\hat{m}_s^{(1)}(\widetilde{X}_i,\beta) - m_s^{(1)}(\widetilde{X}_i,\beta))^2|\widetilde{X}_i] \le c(mh^p b_m^2)^{-1} + ch^{2k}b_m^{-2} + cI[F(\widetilde{X}_i) < 2b_m]$.

This proof is similar to that of Lemma 1 of Xue [5]. Here we omit the details.

Lemma 3.2 *Under Condition* C, *if* β *is the true value of the parameter, we have*

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\widehat{W}_i(\beta) \xrightarrow{L} N(0, \Sigma(\beta)).$$

Proof

$$\begin{aligned}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\widehat{W}_i(\beta) &= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\hat{m}^{(1)}(\widetilde{X}_i,\beta)\big(Y_{i\widehat{G}} - \hat{m}(\widetilde{X}_i,\beta)\big)\\ &= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}m^{(1)}(\widetilde{X}_i,\beta)\big[Y_{iG} - m(\widetilde{X}_i,\beta)\big]\end{aligned}$$

$$
\begin{aligned}
&+ \frac{1}{\sqrt{n}} \sum_{i=1}^{n} m^{(1)}(\widetilde{X}_i, \beta)[Y_{i\widehat{G}} - Y_{iG}] \\
&+ \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \hat{m}^{(1)}(\widetilde{X}_i, \beta)\bigl[m(\widetilde{X}_i, \beta) - \hat{m}(\widetilde{X}_i, \beta)\bigr] \\
&+ \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \bigl[Y_{i\widehat{G}} - m(\widetilde{X}_i, \beta)\bigr]\bigl[\hat{m}^{(1)}(\widetilde{X}_i, \beta) - m^{(1)}(\widetilde{X}_i, \beta)\bigr] \\
=:{}& M_1 + M_2 + M_3 + M_4.
\end{aligned} \tag{3.1}
$$

Employing C_9-C_{11}, similar to the proof of Lemma 1 in Lai *et al.* [11], we have

$$
\begin{aligned}
M_2 &= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} m^{(1)}(\widetilde{X}_i, \beta)(Y_{i\widehat{G}} - Y_{iG}) \\
&= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} m^{(1)}(\widetilde{X}_i, \beta) \frac{m^{(1)}(\widetilde{X}_i, \beta)\delta_i Z_i}{1 - \widehat{G}(Z_i-)} \int_{t<Z_i} \frac{1 - \widehat{G}(t-)}{1 - G(t)} \frac{\sum_{j=1}^{n} dM_j(t)}{Y_n(t)} \\
&= \frac{1}{\sqrt{n}} \sum_{j=1}^{n} \int_{-\infty}^{\tau_F} \left\{ \sum_{i=1}^{N} \frac{m^{(1)}(\widetilde{X}_i, \beta)\delta_i Z_i}{1 - \widehat{G}(Z_i-)} I(Z_i > t) \right\} \frac{1 - \widehat{G}(t-)}{1 - G(t)} \frac{dM_j(t)}{Y_n(t)} + o_p(1) \\
&= -\frac{1}{\sqrt{n}} \sum_{j=1}^{n} \int_{-\infty}^{\tau_F} \left\{ \int_{s>t} s \, d\Gamma_1(s) \right\} \frac{dM_j(t)}{(1 - G(t))(1 - F(t))} + o_p(1) \\
&= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \int_{-\infty}^{\tau_F} E\bigl[m^{(1)}(\widetilde{X}_i, \beta) Y_{iG} I[s < Z_i]\bigr] \frac{dM_j(t)}{(1 - G(t))(1 - F(t))} + o_p(1) \\
&=: M_5 + o_p(1),
\end{aligned}
$$

where $M_i(t) = (1 - \delta_i)I(Z_i < t) - \int_{-\infty}^{t} I(c_i \geq s, y_i > s)\, d\Lambda(s)$, $Y_n(t) = \sum_{i=1}^{n} I(Z_i < t)$. The first item M_5 is a martingale sequence which has a limiting normal distribution with mean 0 and covariance Σ_1 according to the Rebolledo central limit theorem of martingales. We have

$$
\begin{aligned}
M_3 &= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \hat{m}^{(1)}(\widetilde{X}_i, \beta)\bigl[m(\widetilde{X}_i, \beta) - \hat{m}(\widetilde{X}_i, \beta)\bigr] \\
&= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} m^{(1)}(\widetilde{X}_i, \beta)\bigl[m(\widetilde{X}_i, \beta) - \hat{m}(\widetilde{X}_i, \beta)\bigr] \\
&\quad + \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \bigl[\hat{m}^{(1)}(\widetilde{X}_i, \beta) - m^{(1)}(\widetilde{X}_i, \beta)\bigr]\bigl[m(\widetilde{X}_i, \beta) - \hat{m}(\widetilde{X}_i, \beta)\bigr] \\
&=: M_{31} + M_{32}.
\end{aligned}
$$

Write $m_b(\tilde{x}, \beta) = m(\tilde{x}, \beta) f(\tilde{x}) f_b^{-1}(\tilde{x})$, $f_b(\tilde{x}) = \max\{f(\tilde{x}), b_m\}$, then

$$
\begin{aligned}
M_{31} &= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} m^{(1)}(\widetilde{X}_i, \beta)\bigl[m(\widetilde{X}_i, \beta) - \hat{m}(\widetilde{X}_i, \beta)\bigr] \\
&= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} m^{(1)}(\widetilde{X}_i, \beta)\bigl[m_b(\widetilde{X}_i, \beta) - \hat{m}(\widetilde{X}_i, \beta)\bigr]
\end{aligned}
$$

$$
\begin{aligned}
&+ \frac{1}{\sqrt{n}} \sum_{i=1}^{n} m^{(1)}(\widetilde{X}_i, \beta)\big[m(\widetilde{X}_i, \beta) - m_b(\widetilde{X}_i, \beta)\big] \\
&=: M_{311} + M_{312}.
\end{aligned}
$$

For $\forall \varepsilon > 0$, we have

$$
\begin{aligned}
P\big(|M_{312}| > \varepsilon\big) &\le P\left\{ \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \big|m^{(1)}(\widetilde{X}_i, \beta)\big| \big|m(\widetilde{X}_i, \beta) - m_b(\widetilde{X}_i, \beta)\big| > \varepsilon \right\} \\
&= P\left\{ \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \big|m^{(1)}(\widetilde{X}_i, \beta)\big| \big|m(\widetilde{X}_i, \beta)\big| \left| \frac{f_b(\widetilde{X}_i) - f(\widetilde{X}_i)}{f_b(\widetilde{X}_i)} \right| > \varepsilon \right\} \\
&\le P\left\{ \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \big|m^{(1)}(\widetilde{X}_i, \beta)\big| \big|m(\widetilde{X}_i, \beta)\big| I\big[f(\widetilde{X}_i) < b_m\big] > \varepsilon \right\} \\
&\le \frac{1}{\varepsilon} \sqrt{n} P\big(f(\widetilde{X}_i) < b_m\big) \to 0.
\end{aligned}
$$

Hence, $M_{312} = o_p(1)$. Write

$$
\begin{aligned}
\zeta_m(x) &= \frac{1}{mh^p} \sum_{j=n+1}^{n+m} \big[m(\widetilde{X}_j, \beta) - g(\widetilde{X}_j, \beta)\big] K\left(\frac{\widetilde{X}_j - x}{h}\right), \\
\xi_m(x) &= \frac{1}{mh^p} \sum_{j=n+1}^{n+m} \big[m(x, \beta) - m(\widetilde{X}_j, \beta)\big] K\left(\frac{\widetilde{X}_j - x}{h}\right), \\
\phi_m(x) &= \big[f(x)\hat{f}_{b_m}(x) - f_b(x)\hat{f}_m(x)\big] f_b^{-2}(x), \\
\Delta_m(x) &= f_{b_m}(x) - f_b(x).
\end{aligned}
$$

A series simple calculation yields

$$
\begin{aligned}
M_{311} &= \frac{1}{n} \sum_{i=1}^{n} m^{(1)}(\widetilde{X}_i, \beta)\zeta_m(\widetilde{X}_i) f_b^{-1}(\widetilde{X}_i) \\
&\quad + \frac{1}{n} \sum_{i=1}^{n} m^{(1)}(\widetilde{X}_i, \beta)\xi_m(\widetilde{X}_i) f_b^{-1}(\widetilde{X}_i) \\
&\quad + \frac{1}{n} \sum_{i=1}^{n} m^{(1)}(\widetilde{X}_i, \beta) m(\widetilde{X}_i, \beta)\phi_m(\widetilde{X}_i) \\
&\quad + \frac{1}{n} \sum_{i=1}^{n} m^{(1)}(\widetilde{X}_i, \beta)\big[\hat{m}(\widetilde{X}_i, \beta)\hat{f}_{b_m}(\widetilde{X}_i) - m(\widetilde{X}_i, \beta) f(\widetilde{X}_i)\big]\Delta_m(\widetilde{X}_i) f_b^{-2}(\widetilde{X}_i) \\
&\quad + \frac{1}{n} \sum_{i=1}^{n} m^{(1)}(\widetilde{X}_i, \beta)\hat{m}(\widetilde{X}_i, \beta)\Delta_m^2(\widetilde{X}_i) f_b^{-2}(\widetilde{X}_i) \\
&=: \sqrt{n} \sum_{i=1}^{5} J_{mi}, \\
J_{m1} &= \frac{1}{n} \sum_{i=1}^{n} m^{(1)}(\widetilde{X}_i, \beta)\zeta_m(\widetilde{X}_i) f_b^{-1}(\widetilde{X}_i)
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{1}{mh^p}\sum_{j=n+1}^{n+m}\bigl[m(\widetilde{X}_j,\beta)-g(X_j,\beta)\bigr]\int m^{(1)}(x,\beta)K\left(\frac{\widetilde{X}_j-x}{h}\right)f_b^{-1}(x)f(x)\,dx \\
&\quad + \frac{1}{mh^p}\sum_{j=n+1}^{n+m}\bigl[m(\widetilde{X}_j,\beta)-g(X_j,\beta)\bigr] \\
&\quad \times\left[\frac{1}{n}\sum_{i=1}^{n}\frac{m^{(1)}(\widetilde{X}_i,\beta)K(\frac{\widetilde{X}_j-\widetilde{X}_i}{h})}{f_b(\widetilde{X}_i)}-\int\frac{m^{(1)}(x,\beta)K(\frac{\widetilde{X}_j-x}{h})}{f_b(x)}f(x)\,dx\right] \\
&=: J_{m11}+J_{m12},
\end{aligned}
$$

$$
\begin{aligned}
J_{m11} &= \frac{1}{mh^p}\sum_{j=n+1}^{n+m}\bigl[m(\widetilde{X}_j,\beta)-g(X_j,\beta)\bigr]\int m^{(1)}(x,\beta)K\left(\frac{\widetilde{X}_j-x}{h}\right)dx \\
&\quad + \frac{1}{mh^p}\sum_{j=n+1}^{n+m}\bigl[m(\widetilde{X}_j,\beta)-g(X_j,\beta)\bigr]\int m^{(1)}(x,\beta)K\left(\frac{\widetilde{X}_j-x}{h}\right)\bigl[f(x)-f_b(x)\bigr]f_b^{-1}(x)\,dx \\
&=: K_{m1}+K_{m2}.
\end{aligned}
$$

By conditions C_4, C_5, and $\sqrt{m}h^{2k}\to 0$, applying a Taylor expansion, we can prove

$$
\begin{aligned}
K_{m1} &= \frac{1}{mh^p}\sum_{j=n+1}^{n+m}\bigl[m(\widetilde{X}_j,\beta)-g(X_j,\beta)\bigr]\int m^{(1)}(\widetilde{X}_j+\mu h,\beta)K(\mu)h^p\,d\mu \\
&= \frac{1}{m}\sum_{j=n+1}^{n+m}\bigl[m(\widetilde{X}_j,\beta)-g(X_j,\beta)\bigr]\int m^{(1)}(\widetilde{X}_j,\beta)K(\mu)h^p\,d\mu + o_p\bigl(m^{-\frac{1}{2}}\bigr) \\
&= \frac{1}{m}\sum_{j=n+1}^{n+m}\bigl[m(\widetilde{X}_j,\beta)-g(X_j,\beta)\bigr]m^{(1)}(\widetilde{X}_j,\beta) + o_p\bigl(m^{-\frac{1}{2}}\bigr).
\end{aligned}
$$

Notice that $f(\tilde{x})-f_b(\tilde{x})=0$ when $f(\tilde{x})\ge b_m$. By conditions C_2-C_4, it is easy to see that $E(\sqrt{m}K_{m2})\to 0$. Hence, we have $K_{m2}=o_p(m^{-\frac{1}{2}})$.

Using the standard kernel estimation method, we can prove that $J_{m12}=o_p(m^{-\frac{1}{2}})$.

Now, let us prove $J_{mi}=o_p(m^{-\frac{1}{2}})$, $i=2,\ldots,5$. By conditions C_4 and C_5, applying a Taylor expansion, we have

$$
\begin{aligned}
E\bigl[\xi_m^2(\widetilde{X}_i)|\widetilde{X}_i\bigr] &= \operatorname{Var}\bigl[\xi_m(\widetilde{X}_i)|\widetilde{X}_i\bigr]+\bigl[E\bigl[\xi_m(\widetilde{X}_i)|\widetilde{X}_i\bigr]\bigr]^2 \\
&\le \frac{1}{mh^p}\int\bigl[m(\widetilde{X}_i,\beta)-m(\widetilde{X}_i+h\mu,\beta)\bigr]^2K^2(\mu)\,d\mu \\
&\quad + \left[\int\bigl[m(\widetilde{X}_i,\beta)-m(\widetilde{X}_i+h\mu,\beta)\bigr]K(\mu)\,d\mu\right]^2 \\
&\le c\bigl(mh^p\bigr)+ch^{2k}.
\end{aligned}
\tag{3.2}
$$

According to the condition independence, together with (3.2), it is easily shown that $E(\sqrt{m}J_{m2})^2\to 0$, hence $J_{m2}=o_p(m^{-\frac{1}{2}})$. Now

$$
\begin{aligned}
J_{m3} &= \frac{1}{n}\sum_{i=1}^{n}m^{(1)}(\widetilde{X}_i,\beta)m(\widetilde{X}_i,\beta)\phi_m(\widetilde{X}_i) \\
&= \frac{1}{n}\sum_{i=1}^{n}m^{(1)}(\widetilde{X}_i,\beta)m(\widetilde{X}_i,\beta)\phi_m(\widetilde{X}_i)I\bigl[f(\widetilde{X}_i)<b_m,\hat{f}_m(\widetilde{X}_i)<b_m\bigr]
\end{aligned}
$$

$$\begin{aligned}
&+\frac{1}{n}\sum_{i=1}^{n} m^{(1)}(\widetilde{X}_i,\beta)m(\widetilde{X}_i,\beta)\phi_m(\widetilde{X}_i)I\big[f(\widetilde{X}_i)\geq b_m,\hat{f}_m(\widetilde{X}_i)<b_m\big]\\
&+\frac{1}{n}\sum_{i=1}^{n} m^{(1)}(\widetilde{X}_i,\beta)m(\widetilde{X}_i,\beta)\phi_m(\widetilde{X}_i)I\big[f(\widetilde{X}_i)<b_m,\hat{f}_m(\widetilde{X}_i)\geq b_m\big]\\
&=: J_{m31}+J_{m32}+J_{m33},
\end{aligned}$$

$$\begin{aligned}
J_{m31} &= \frac{1}{n}\sum_{i=1}^{n} m^{(1)}(\widetilde{X}_i,\beta)m(\widetilde{X}_i,\beta)b_m^{-1}\big[f(\widetilde{X}_i)-\hat{f}_m(\widetilde{X}_i)\big]I\big[f(\widetilde{X}_i)<b_m,-b_m<\hat{f}_m(\widetilde{X}_i)<b_m\big]\\
&\quad+\frac{1}{n}\sum_{i=1}^{n} m^{(1)}(\widetilde{X}_i,\beta)m(\widetilde{X}_i,\beta)b_m^{-1}\big[f(\widetilde{X}_i)-\hat{f}_m(\widetilde{X}_i)\big]I\big[f(\widetilde{X}_i)<b_m,\hat{f}_m(\widetilde{X}_i)<-b_m\big]\\
&=: L_{m1}+L_{m2}.
\end{aligned}$$

By condition C_4, we easily get

$$\sup_{\tilde{x}}\big|\hat{f}_m(\tilde{x})-f(\tilde{x})\big| = O_p\big[\big(mh^p\big)^{-\frac{1}{2}}\big]+O_p\big(h^k\big). \tag{3.3}$$

Together with condition C_5 and (3.3), and the Markov inequality, it is easily proved that $L_{m1}=o_p(m^{-\frac{1}{2}})$. In addition, for $\forall\varepsilon>0$, we have

$$P\big(\sqrt{m}|L_{m2}|>\varepsilon\big)\leq P\Big(\sup_{\tilde{x}}\big|\hat{f}_m(\tilde{x})-f(\tilde{x})\big|>b_m\Big)\to 0.$$

Hence, we have $J_{m31}=o_p(m^{-\frac{1}{2}})$. By an argument similar to J_{m31}, we obtain

$$\sqrt{m}|J_{m32}|\to 0,\qquad \sqrt{m}|J_{m33}|\to 0.$$

Hence, we have $J_{m3}=o_p(m^{-\frac{1}{2}})$. Now

$$\begin{aligned}
J_{m4} &= \frac{1}{n}\sum_{i=1}^{n} m^{(1)}(\widetilde{X}_i,\beta)\big[\hat{m}(\widetilde{X}_i,\beta)-m(\widetilde{X}_i,\beta)\big]\Delta_m^2(\widetilde{X}_i)f_b^{-2}(\widetilde{X}_i)\\
&\quad+\frac{1}{n}\sum_{i=1}^{n} m^{(1)}(\widetilde{X}_i,\beta)\big[\hat{m}(\widetilde{X}_i,\beta)-m(\widetilde{X}_i,\beta)\big]\Delta_m(\widetilde{X}_i)f_b^{-1}(\widetilde{X}_i)\\
&\quad+\frac{1}{n}\sum_{i=1}^{n} m^{(1)}(\widetilde{X}_i,\beta)m(\widetilde{X}_i,\beta)\Delta_m^2(\widetilde{X}_i)f_b^{-2}(\widetilde{X}_i)\\
&\quad+\frac{1}{n}\sum_{i=1}^{n} m^{(1)}(\widetilde{X}_i,\beta)m(\widetilde{X}_i,\beta)\Delta_m(\widetilde{X}_i)\big[b_m-f(\widetilde{X}_i)\big]b_m^{-2}I\big[f(\widetilde{X}_i)<b_m\big].
\end{aligned}$$

Notice that

$$\sup_{\tilde{x}}\big|\Delta_m(\tilde{x})\big|\leq\sup_{\tilde{x}}\big|\hat{f}_m(x)-f(\tilde{x})\big| = O_p\big[\big(mh^p\big)^{-\frac{1}{2}}\big]+O_p\big(h^k\big). \tag{3.4}$$

Together with $\sqrt{m}P(f(\widetilde{X}_i)<b_m)\to 0$, applying conditional independence, we can obtain $J_{m4}=o_p(m^{-\frac{1}{2}})$.

Similarly, by conditional independence and (3.4), we can prove that $J_{m5} = o_p(m^{-\frac{1}{2}})$. Hence, we have

$$M_{31} = \frac{\sqrt{n}}{m}\sum_{j=n+1}^{n+m}\big[m(\widetilde{X}_j,\beta) - g(X_j,\beta)\big]m^{(1)}(\widetilde{X}_j,\beta) + o_p(1).$$

Consider the sth ($s = 1,\ldots,p$) component of M_{32}, from the Cauchy-Schwarz inequality and Lemma 3.1, we have

$$\begin{aligned}
E|M_{32s}| &= \frac{1}{\sqrt{n}}E\left|\sum_{i=1}^{n}\big[\hat{m}_s^{(1)}(\widetilde{X}_i,\beta) - m_s^{(1)}(\widetilde{X}_i,\beta)\big]\big[m(\widetilde{X}_i,\beta) - \hat{m}(\widetilde{X}_i,\beta)\big]\right| \\
&\leq \frac{1}{\sqrt{n}}\sum_{i=1}^{n}E\big|\big[\hat{m}_s^{(1)}(\widetilde{X}_i,\beta) - m_s^{(1)}(\widetilde{X}_i,\beta)\big]\big[m(\widetilde{X}_i,\beta) - \hat{m}(\widetilde{X}_i,\beta)\big]\big| \\
&\leq \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\big\{E\big[\hat{m}_s^{(1)}(\widetilde{X}_i,\beta) - m_s^{(1)}(\widetilde{X}_i,\beta)\big]^2 E\big[m(\widetilde{X}_i,\beta) - \hat{m}(\widetilde{X}_i,\beta)\big]^2\big\}^{\frac{1}{2}} \\
&= o(1).
\end{aligned}$$

Hence, we have $M_{32} = o_p(1)$,

$$\begin{aligned}
M_3 &= \frac{\sqrt{n}}{m}\sum_{j=n+1}^{n+m}\big[m(\widetilde{X}_j,\beta) - g(X_j,\beta)\big]m^{(1)}(\widetilde{X}_j,\beta) + o_p\big(m^{-\frac{1}{2}}\big) =: M_6 + o_p(1), \\
M_4 &= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\big[Y_{i\widehat{G}} - m(\widetilde{X}_i,\beta)\big]\big[\hat{m}^{(1)}(\widetilde{X}_i,\beta) - m^{(1)}(\widetilde{X}_i,\beta)\big] \\
&= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\big[Y_{iG} - m(\widetilde{X}_i,\beta)\big]\big[\hat{m}^{(1)}(\widetilde{X}_i,\beta) - m^{(1)}(\widetilde{X}_i,\beta)\big] \\
&\quad + \frac{1}{\sqrt{n}}\sum_{i=1}^{n}[Y_{i\widehat{G}} - Y_{iG}]\big[\hat{m}^{(1)}(\widetilde{X}_i,\beta) - m^{(1)}(\widetilde{X}_i,\beta)\big] \\
&=: M_{41} + M_{42}.
\end{aligned}$$

From Wang [3, 4], we can obtain $M_{42} = o_p(1)$.

Write $U = \{X_j,\widetilde{X}_j\}_{j=n+1}^{n+m}$, consider the sth ($s = 1,\ldots,p$) component of M_{41}, from conditional independence and Lemma 3.1 we have

$$\begin{aligned}
E(M_{41s})^2 &= \frac{1}{n}E\left\{\sum_{i=1}^{n}\big[Y_{iG} - m(\widetilde{X}_i,\beta)\big]\big[\hat{m}_s^{(1)}(\widetilde{X}_i,\beta) - m_s^{(1)}(\widetilde{X}_i,\beta)\big]\right\}^2 \\
&= \frac{1}{n}E\left\{E\left\{\sum_{i=1}^{n}\eta_i\big[\hat{m}_s^{(1)}(\widetilde{X}_i,\beta) - m_s^{(1)}(\widetilde{X}_i,\beta)\big]^2\middle|U\right\}\right\} \\
&= \frac{1}{n}\sum_{i=1}^{n}E\big\{\eta_i\big[\hat{m}_s^{(1)}(\widetilde{X}_i,\beta) - m_s^{(1)}(\widetilde{X}_i,\beta)\big]\big\}^2 \\
&= o(1).
\end{aligned}$$

So $M_{41} = o_p(1)$, then $M_4 = o_p(1)$. Therefore, we have

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\widehat{W}_i(\beta) = M_1 + M_5 + M_6 + o_p(1). \tag{3.5}$$

From the central limit theorem and $\frac{n}{m} \to \gamma$, we have

$$M_1 \xrightarrow{L} N\bigl(0, \Sigma_0(\beta)\bigr), \qquad M_5 \xrightarrow{L} N\bigl(0, \Sigma_1(\beta)\bigr), \tag{3.6}$$

$$M_6 \xrightarrow{L} N\bigl(0, \Sigma_2(\beta)\bigr). \tag{3.7}$$

In addition, M_1 and M_6 are independent of each other, M_5 and M_6 are independent of each other. Then a simple calculation yields

$$EM_1M_6 = 0, \qquad EM_5M_6 = 0, \qquad EM_1M_5 \xrightarrow{P} -\Sigma_1(\beta). \tag{3.8}$$

From the central limit theorem and by (3.6)-(3.8), Lemma 3.2 is proved. □

Lemma 3.3 *Under Condition* C, *if β is the true value of the parameter, we have*

$$\frac{1}{n}\sum_{i=1}^{n}\widehat{W}_i(\beta)\widehat{W}_i^{\tau}(\beta) \xrightarrow{P} \Sigma_0(\beta).$$

Proof After a complex calculation, we have

$$\frac{1}{n}\sum_{i=1}^{n}\widehat{W}_i(\beta)\widehat{W}_i^{\tau}(\beta) = \frac{1}{n}\sum_{i=1}^{n}m^{(1)}(\widetilde{X}_i,\beta)\bigl(m^{(1)}(\widetilde{X}_i,\beta)\bigr)^{\tau}\eta_i^2 + o_p(1).$$

By the law of large numbers, we obtain Lemma 3.3. □

Lemma 3.4 *Let $Z \xrightarrow{L} N(0, I_p)$, where I_p is the $p \times p$ identity matrix. Let Q be a $p \times p$ nonnegative definite matrix with eigenvalues $\omega_1, \ldots, \omega_p$. Then it follows that*

$$Z^{\tau}QZ \xrightarrow{L} \omega_1\chi_{1,1}^2 + \omega_2\chi_{2,1}^2 + \cdots + \omega_p\chi_{p,1}^2.$$

Proof of Theorem 2.1 We have

$$\hat{\beta} = \arg\min_{\beta}\frac{1}{2}\sum_{i=1}^{n}\bigl[Y_{i\widehat{G}} - \hat{m}(\widetilde{X}_i,\beta)\bigr]^2.$$

Let

$$L(\beta) = \frac{1}{2}\sum_{i=1}^{n}\bigl[Y_{i\widehat{G}} - \hat{m}(\widetilde{X}_i,\beta)\bigr]^2.$$

Using the Lagrange multiplier method, let

$$\frac{\partial L(\beta)}{\partial\beta} = \left(\frac{\partial L(\beta)}{\partial\beta_1}, \ldots, \frac{\partial L(\beta)}{\partial\beta_p}\right) = 0,$$

that is,

$$\sum_{i=1}^{n} \hat{m}^{(1)}(\widetilde{X}_i, \beta)\big[Y_{i\widehat{G}} - \hat{m}(\widetilde{X}_i, \beta)\big] = 0.$$

Then we have

$$\frac{1}{n}\sum_{i=1}^{n} \hat{m}^{(1)}(\widetilde{X}_i, \hat{\beta})\big[Y_{i\widehat{G}} - \hat{m}(\widetilde{X}_i, \beta)\big] = 0. \tag{3.9}$$

Applying the Taylor expansion to $\hat{m}(\widetilde{X}_i, \beta)$ and $\hat{m}^{(1)}(\widetilde{X}_i, \beta)$ in (3.9), we can obtain

$$\begin{aligned}
&\hat{m}^{(1)}(\widetilde{X}_i, \hat{\beta}) = \hat{m}^{(1)}(\widetilde{X}_i, \beta) + o_p(1), \\
&\hat{m}(\widetilde{X}_i, \hat{\beta}) = \hat{m}(\widetilde{X}_i, \beta) + \hat{m}^{(1)}\big(\widetilde{X}_i, \beta + \theta(\hat{\beta} - \beta)\big)(\hat{\beta} - \beta) + o_p(1), \\
&0 = \frac{1}{n}\sum_{i=1}^{n}\big[\hat{m}^{(1)}(\widetilde{X}_i, \beta) + o_p\big(n^{\frac{1}{2}}\big)\big] \\
&\quad \times \big[Y_{i\widehat{G}} - \hat{m}(\widetilde{X}_i, \beta) - \hat{m}^{(1)}\big(\widetilde{X}_i, \beta + \theta(\hat{\beta} - \beta)\big)(\hat{\beta} - \beta) + o_p\big(n^{-\frac{1}{2}}\big)\big] \\
&\quad = \frac{1}{n}\sum_{i=1}^{n} \hat{m}^{(1)}(\widetilde{X}_i, \beta)\big[Y_{i\widehat{G}} - \hat{m}(\widetilde{X}_i, \beta)\big] \\
&\qquad - \frac{1}{n}\sum_{i=1}^{n} \hat{m}^{(1)}(\widetilde{X}_i, \beta)\hat{m}^{(1)}\big(\widetilde{X}_i, \beta + \theta(\hat{\beta} - \beta)\big)(\hat{\beta} - \beta) + o_p\big(n^{-\frac{1}{2}}\big).
\end{aligned}$$

So

$$\hat{\beta} - \beta = \widehat{\Xi}^{-1}(\beta)\left\{\frac{1}{n}\sum_{i=1}^{n} \hat{m}^{(1)}(\widetilde{X}_i, \beta)\big[Y_{i\widehat{G}} - \hat{m}(\widetilde{X}_i, \beta)\big]\right\} + O_p\big(n^{-\frac{1}{2}}\big),$$

where $\widehat{\Xi}^{-1}(\beta) = \frac{1}{n}\sum_{i=1}^{n} \hat{m}^{(1)}(\widetilde{X}_i, \beta)\hat{m}^{(1)}(\widetilde{X}_i, \beta + \theta(\hat{\beta} - \beta))(\hat{\beta} - \beta)$, $\theta \in (0,1)$ is a constant. This, together with Lemma 3.2, easily proves that $\widehat{\Xi}^{-1}(\beta) \xrightarrow{P} \Xi(\beta)$. That is,

$$\hat{\beta} - \beta = \frac{1}{n}\widehat{\Xi}^{-1}(\beta)\sum_{i=1}^{n} \widehat{W}_i(\beta) + O_p\big(n^{-\frac{1}{2}}\big).$$

From Lemmas 3.2, 3.3, and 3.4, we can obtain

$$\sqrt{n}(\hat{\beta} - \beta) \xrightarrow{L} N\big(0, \Xi^{-1}\Sigma(\beta)\Xi^{-1}\big).$$

Theorem 2.1 is proved. □

Proof of Theorem 2.2 Applying a Taylor expansion to equation (2.3), we have

$$\begin{aligned}
\hat{l}(\beta) &= 2\sum_{i=1}^{n} \log\big(1 + \lambda^{\tau}\widehat{W}_i(\beta)\big) \\
&= 2\sum_{i=1}^{n}\left[\log 1 + \big(\log' x\big)_{x=1}\big(\lambda^{T}\widehat{W}_i\big) + \frac{(\log x)^{(2)}|_{x=1}}{2}\big(\lambda^{\tau}\widehat{W}_i\big)^2 + \frac{(\log x)^{(3)}|_{x=\xi}}{3}\big(\lambda^{\tau}\widehat{W}_i\big)^3\right]
\end{aligned}$$

$$\begin{aligned} &(0 < \xi < 1) \\ &= 2\sum_{i=1}^{n}\left[\lambda^{\tau}\widehat{W}_i - \frac{1}{2}\left(\lambda^{\tau}\widehat{W}_i\right)^2\right] + o_p(1). \end{aligned} \tag{3.10}$$

By Lemmas 3.2 and 3.3, using the result of Wang [2–4], we can prove

$$\begin{cases} \max_{1\le i\le n}\widehat{W}_i(\beta) = o_p(n^{\frac{1}{2}}), \\ \lambda = O_p(n^{-\frac{1}{2}}), \\ \frac{1}{n}\sum_{i=1}^{n}\widehat{W}_i(\beta)\widehat{W}_i^{\tau}(\beta) = o_p(1). \end{cases} \tag{3.11}$$

Again,

$$\begin{aligned} 0 &= \frac{1}{n}\sum_{i=1}^{n}\frac{\widehat{W}_i(\beta)}{1+\lambda^{\tau}\widehat{W}_i(\beta)} \\ &= \frac{1}{n}\sum_{i=1}^{n}\frac{\widehat{W}_i(\beta)(1+\lambda^{\tau}\widehat{W}_i(\beta)) - \widehat{W}_i(\beta)\lambda^{\tau}\widehat{W}_i(\beta)}{1+\lambda^{\tau}\widehat{W}_i(\beta)} \\ &= \frac{1}{n}\sum_{i=1}^{n}\widehat{W}_i(\beta) - \frac{1}{n}\sum_{i=1}^{n}\lambda^{\tau}\left(\widehat{W}_i(\beta)\right)^2 + \frac{1}{n}\sum_{i=1}^{n}\frac{\widehat{W}_i(\beta)\lambda^{\tau}\lambda\widehat{W}_i(\beta)\widehat{W}_i^{\tau}(\beta)}{1+\lambda^{\tau}\widehat{W}_i(\beta)}. \end{aligned} \tag{3.12}$$

By (3.11), we can obtain

$$\frac{1}{n}\sum_{i=1}^{n}\lambda^{\tau}\widehat{W}_i(\beta) = \frac{1}{n}\sum_{i=1}^{n}\lambda^{\tau}\left(\lambda^{\tau}\widehat{W}_i(\beta)\right)^2 + o_p(1). \tag{3.13}$$

By (3.13), we can obtain

$$\lambda = \left(\sum_{i=1}^{n}\widehat{W}_i(\beta)\widehat{W}_i^{\tau}(\beta)\right)^{-1}\left(\sum_{i=1}^{n}\widehat{W}_i(\beta)\right) + o_p\left(n^{-\frac{1}{2}}\right). \tag{3.14}$$

By (3.10), (3.13), (3.14), we can obtain

$$\begin{aligned} \hat{l}(\beta) &= \lambda^{\tau}\left(\sum_{i=1}^{n}\widehat{W}_i(\beta)\right) + o_p(1) \\ &= \left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\widehat{W}_i(\beta)\right)^{\tau}\widehat{\Sigma}_0^{-1}(\beta)\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\widehat{W}_i(\beta)\right) + o_p(1). \end{aligned} \tag{3.15}$$

By (3.15) and Lemmas 3.2, 3.3, and 3.4, we can obtain

$$\hat{l}(\beta) = \left\{\frac{1}{\sqrt{n}}\Sigma^{-\frac{1}{2}}(\beta)\sum_{i=1}^{n}\widehat{W}_i(\beta)\right\}^{\tau}D(\beta)\left\{\frac{1}{\sqrt{n}}\Sigma^{-\frac{1}{2}}(\beta)\sum_{i=1}^{n}\widehat{W}_i(\beta)\right\} + o_p(1),$$

where $D(\beta) = \Sigma^{\frac{1}{2}}(\beta)\Sigma_0^{-1}(\beta)\Sigma^{\frac{1}{2}}(\beta)$.

Write $\widetilde{D} = \operatorname{diag}(\omega_1,\ldots,\omega_p)$ where ω_i $(i = 1,\ldots,p)$ are the eigenvalues of D, then there exists an orthogonal matrix Q which makes $Q^{\tau}\widetilde{D}Q = D(\beta)$. Therefore, we have

$$\hat{l}(\beta) = \left\{\frac{1}{\sqrt{n}}Q\Sigma^{-\frac{1}{2}}(\beta)\sum_{i=1}^{n}\widehat{W}_i(\beta)\right\}^{\tau}\widetilde{D}(\beta)\left\{\frac{1}{\sqrt{n}}Q\Sigma^{-\frac{1}{2}}(\beta)\sum_{i=1}^{n}\widehat{W}_i(\beta)\right\} + o_p(1). \tag{3.16}$$

From Lemma 3.2, we have

$$\frac{1}{\sqrt{n}} Q \Sigma^{-\frac{1}{2}}(\beta) \sum_{i=1}^{n} \widehat{W}_i(\beta) \xrightarrow{L} N(0, I_p). \tag{3.17}$$

By (3.16) and (3.17), we finish the proof of Theorem 2.2. □

Proof of Theorem 2.3 Review the definition of $\hat{l}_{\mathrm{ad}}(\beta)$, combined with (3.16), we can infer

$$\hat{l}_{\mathrm{ad}}(\beta) = \left\{ \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \widehat{W}_i(\beta) \right\}^{\tau} \widehat{\Sigma}^{-1}(\beta) \left\{ \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \widehat{W}_i(\beta) \right\} + o_p(1).$$

Combined with the process of the above proof, similar to the proof of Lemma 3.2, it is easily shown that $\widehat{\Sigma}(\beta) \xrightarrow{P} \Sigma(\beta)$.

By Lemmas 3.2, 3.4, and (3.17), we can prove

$$\hat{l}_{\mathrm{ad}}(\beta) \xrightarrow{L} \chi_p^2.$$ □

4 Conclusions

In this text, with the help of validation data, the empirical likelihood method is extended to the nonlinear error-in-variables regression models with randomly censored response. We construct an estimated empirical log-likelihood function $\hat{l}(\beta)$ of the unknown interesting parameter β, and we get the asymptotic distribution of $\hat{l}(\beta)$. By giving estimators of the weights ω_i and Monte Carlo simulation method, we construct the confidence region for the parameter β. To avoid estimating ω_i, the adjusted empirical log-likelihood $\hat{l}_{\mathrm{ad}}(\beta)$ has been defined and the asymptotic distribution of $\hat{l}_{\mathrm{ad}}(\beta)$ is obtained. Using the result, we may better construct a confidence region for the parameter β. The theoretical analysis shows that the empirical likelihood method can be applied to the nonlinear model with complex incomplete data, and it can be used to obtain a better result.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

All authors contributed equally to the writing of this paper. All authors read and approved the final manuscript.

Acknowledgements

The authors are grateful to the referee for carefully reading the manuscript and for offering comments which enabled them to improve the paper. The research of Y Wu was supported by the Humanities and Social Sciences Foundation for the Youth Scholars of Ministry of Education of China (12YJCZH217) and the Key NSF of Anhui Educational Committe (KJ2014A255). The research of L Fang was supported the Anhui provincial natural science research project of university foundation (KJ2013B304) and the natural science research project of the Tongling University foundation (2012tlxy13).

Received: 24 August 2015 Accepted: 21 November 2015 Published online: 03 December 2015

References

1. Sepanski, JH, Lee, LF: Semiparametric estimation of nonlinear error-in-variables models with validation study. J. Nonparametr. Stat. **4**, 365-394 (1995)
2. Wang, QH: Estimation of linear error-in-covariables model with validation data under random censorship. J. Multivar. Anal. **74**, 245-266 (1999)
3. Wang, QH, Rao, JNK: Empirical likelihood-based in linear error-in-covariables model with validation data. Biometrika **89**(2), 345-358 (2002)
4. Wang, QH: Survival Data Analysis. Science Press, Beijing (2006)
5. Xue, L: Empirical likelihood inference in nonlinear semiparametric EV models with validation data. Acta Math. Sin. **49**(1), 145-154 (2006)

6. Fang, L, Hu, F: Empirical likelihood dimension reduction inference in nonlinear EV models with validation data. J. Math. **32**(1), 113-120 (2012)
7. Buckley, J, James, I: Linear regression with censored data. Biometrika **66**, 429-436 (1979)
8. Li, G, Wang, QH: Empirical likelihood confidence regression analysis with right censored data. Stat. Sin. **13**(1), 51-68 (2003)
9. Cheng, F, Li, G, Feng, S, Xue, L: Empirical likelihood inference for nonlinear regression model with right censored data. Acta Math. Appl. Sin. **33**(1), 130-140 (2010)
10. Rao, JNK, Scott, AJ: The analysis of categorical data from complex sample surveys: chi-squared test for goodness of fit and independence in two-way tables. J. Am. Stat. Assoc. **76**, 221-230 (1981)
11. Lai, TL, Ying, ZL, Zheng, ZK: Asymptotic normality of a class of adaptive statistic with applications to synthetic data methods for censored regression. J. Multivar. Anal. **52**, 259-279 (1995)