第47卷第9期 中国科学技术大学学报 Vol.47,No.9
2017年9月 JOURNAL OF UNIVERSITY OF SCIENCE AND TECHNOLOGY OF CHINA Sep. 2017

文章编号:0253-2778(2017)9-0755-07

带有右删失的最小乘积相对误差估计

胡大海,张伟平

(中国科学技术大学管理学院统计与金融系,安徽合肥 230026)

摘要:考虑了带有右删失的乘积回归模型. 在随机右删失的情形下,通过逆概率加权的方法将乘积相对误差准则推广到右删失的情形. 在一定的正则条件下,建立了估计的相合性和渐近正态性. 最后,通过数值模拟展示所提出方法的效果.

关键词:乘积回归模型;相对误差;删失;逆概率加权

中图分类号:O212.1 **文献标识码**:A doi:10.3969/j.issn.0253-2778.2017.09.007

2010 Mathematics Subject Classification: 62N02

引用格式:胡大海,张伟平. 带有右删失的最小乘积相对误差估计[J]. 中国科学技术大学学报,2017,47(9):755-761.

HU Dahai, ZHANG Weiping. Least product relative error estimation with right censoring[J]. Journal of University of Science and Technology of China, 2017,47(9):755-761.

Least product relative error estimation with right censoring

HU Dahai, ZHANG Weiping

(*Department of Statistics and Finance, School of Management, University of Science and Technology of China, Hefei* 230026, *China*)

Abstract: The multiplicative regression model with right censoring was considered. Under the random right censorship, the criterion of the least product relative error (LPRE) was extended to the case with right censoring. Under some regular conditions, the consistency and asymptotic normality were established. Finally, some simulations were conducted to examine the finite performance of the proposed method.

Key words: multiplicative regression model; relative error; censoring; inverse probability weighted

0 引言

在应用中,常常遇到正的响应变量,如生物个体的生存时间、机械系统的使用寿命等. 对于处理正的响应变量,我们一般会考虑乘积回归模型(multiplicative regression model),

$$T_i=\exp(X_i^{\mathrm{T}}\beta_0)\varepsilon_i,\ i=1,\cdots,n \tag{1}$$

式中,T_i 是一维的响应变量,X_i 是 p 维的协变量向量,β_0 是未知的回归参数向量,ε_i 是恒正的误差项. 当 T_i 是生存时间时,模型(1)是生存分析中的 AFT 模型(accelerate failure time model).

在很多情形下,对于正的响应变量,应用者更关注相对误差(relative error)而不是绝对误差(absolute error).基于不同的相对目标, Chen 等[1]率先提出了如下两种类型的相对误差:一种是相对于响应变量 T_i,即 $|T_i-\exp(X_i^{\mathrm{T}}\beta)|/T_i$;另一种是

收稿日期:2017-03-09;**修回日期**:2017-06-11

基金项目:国家自然科学基金(11671374)资助.

作者简介:胡大海,男,1988年生,硕士.研究方向:相对误差,测量误差.E-mail: dhh@mail.ustc.edu.cn

通讯作者:张伟平,博士/副教授.E-mail: zwp@ustc.edu.cn

相对于 T_i 的预测值 $\exp(X_i^{\mathrm{T}}\beta)$，即 $|T_i-\exp(X_i^{\mathrm{T}}\beta)|/\exp(X_i^{\mathrm{T}}\beta)$.基于这两种类型的相对误差，Chen 等[1]提出最小相对误差（least absolute relative error，LARE）估计，即最小化如下目标函数：

$$\sum_{i=1}^{n}\left\{\frac{|T_i-\exp(X_i^{\mathrm{T}}\beta)|}{\exp(X_i^{\mathrm{T}}\beta)}+\frac{|T_i-\exp(X_i^{\mathrm{T}}\beta)|}{T_i}\right\}.$$

基于 LARE 准则，Zhang 和 Wang[2]考虑了乘积部分线性模型，Zhang 等[3]考虑了乘积函数型数据. 但是，LARE 估计的目标函数在某些点处不可导，且 LARE 估计的方差含有 ε_i 的密度函数，从而使得方差的估计变得复杂. 基于这些 LARE 准则的不足，Chen 等[4]提出了最小乘积相对误差（least product relative error，LPRE）估计，即最小化

$$\mathrm{LPRE}_n(\beta)=\sum_{i=1}^{n}\left\{\frac{|T_i-\exp(X_i^{\mathrm{T}}\beta)|}{\exp(X_i^{\mathrm{T}}\beta)}\times\frac{|T_i-\exp(X_i^{\mathrm{T}}\beta)|}{T_i}\right\} \tag{2}$$

LPRE 估计的目标函数有如下优势：第一，它是无限可导的；第二，它是严格凸的. 这些优势使得 LPRE 准则被广泛应用. 基于 LPRE 准则，Wang 等[5]考虑了变点的估计；Liu 等[6]考虑了变量选择问题.

虽然 LARE 准则和 LPRE 准则有很多优点，但是它们都是基于完全观测的数据. 然而，在一些领域特别是生存分析中，个体的生存时间被删失的情形很多，比如参加某项医学研究的某个病人在研究结束或者死亡之前退出；或者一项研究在终止之前，仍有部分个体还活着；或者，某个病人可能患很多种疾病，其他疾病造成的死亡也是一种删失.因此，将相对误差准则推广到删失的情形在应用中是十分必要且重要的.

对模型(1)进行对数变换，变成如下线性模型：

$$\ln T_i=X_i^{\mathrm{T}}\beta_0+\varepsilon_i^*,\ i=1,\cdots,n \tag{3}$$

式中，$\varepsilon_i^*=\ln\varepsilon_i$. 当 $E[\varepsilon_i^*]=0$ 时，所有基于最小二乘准则处理响应变量删失的方法都可以应用到模型(3)中，主要有 Miller 估计[7]，Buckley-James 估计[8-10]，KSV 估计[11]，加权最小二乘（WLS）估计[12-15]. 其中，KSV 估计和 WLS 估计不需要迭代，比较易于使用.但是，第一，上述方法都是基于变换后的线性模型(3)，当 $E[\varepsilon_i^*]\neq 0$ 时，基于模型(3)可能导致参数估计的不相合. 第二，变换后的模型分析是基于绝对误差的. 因此，基于相对误差准则考虑非变换的模型(1)显得尤其重要. 针对这一重要的问题，目前仍未有学者考虑过此问题.

基于乘积相对误差（LPRE）准则，本文研究了随机右删失情形下模型(1)的参数估计问题，使得 LPRE 准则在生存分析中可以广泛地使用. 通过逆概率加权的方法，给出了参数估计的方法，建立了估计的相合性和渐近正态性. 大量的数值模拟表明，在有限样本情形下，所提出方法是可行的.

在节 1，我们提出了带有右删失的乘积相对误差估计，并研究了该估计的相合性和渐近正态性；定理的详细证明在节 2 中给出；节 3 通过数值模拟来展示所提方法的有限样本下的效果.

1 模型估计和渐近性质

对个体 i，记 T_i 为失效时间，X_i 是相应的协变量，C_i 是 T_i 的删失时间，其中，$i=1,\cdots,n$. 设 $(T_i,X_i,C_i,\varepsilon_i)$，$i=1,\cdots,n$，是独立同分布的.在完全数据下，我们可以观测到每个个体的失效时间 T_i. 在右删失的情形下，我们只能观测到 $Y_i=\min\{T_i,C_i\}$ 和 $\delta_i=I(T_i\leqslant C_i)$，其中，$I(\cdot)$是一个示性函数.通过一些简单的代数运算，式(2)可以化简为

$$\mathrm{LPRE}_n(\beta)=\sum_{i=1}^{n}\{T_i^{-1}\exp(X_i^{\mathrm{T}}\beta)+T_i\exp(-X_i^{\mathrm{T}}\beta)-2\} \tag{4}$$

最小化 $\mathrm{LPRE}_n(\beta)$等价于求解如下估计方程：

$$\sum_{i=1}^{n}X_i\{T_i^{-1}\exp(X_i^{\mathrm{T}}\beta)-T_i\exp(-X_i^{\mathrm{T}}\beta)\}=0.$$

然而，在随机右删失情形下，失效时间 T_i 不能被观测到，我们只能观测到(Y_i,δ_i).设函数 $S(\cdot)$ 是删失时间 C_i 的生存函数.基于观测到的数据(Y_i,δ_i)，通过逆概率加权办法构造如下估计方程：

$$\sum_{i=1}^{n}\frac{\delta_i}{S(Y_i)}X_i\{Y_i^{-1}\exp(X_i^{\mathrm{T}}\beta)-Y_i\exp(-X_i^{\mathrm{T}}\beta)\}=0 \tag{5}$$

逆概率加权有一个直观的解释. 在随机右删失的情形下，如果第 i 个个体真实的失效时间为 T_i，那么该个体的失效时间将会以概率 $S(T_i)=P(C_i>T_i|T_i)$被观测到. 也就是说，如果失效时间 T_i 被观测到，那么在平均的意义下，将会有 $1/S(T_i)$个个体的真实失效时间也为 T_i，但是我们只观测到了其中的 1 个个体，其他的个体被删失了.

但是，函数 $S(\cdot)$是未知的，基于数据$(Y_i,1-\delta_i)$，$i=1,\cdots,n$，构造 $S(\cdot)$ 的 Kaplan-Meier

估计 $\hat{S}(\cdot)$. 因此,将方程(5)中的 $S(\cdot)$ 替换成 $\hat{S}(\cdot)$，得到最终的估计方程：

$$\sum_{i=1}^{n}\frac{\delta_i}{\hat{S}(Y_i)}X_i\{Y_i^{-1}\exp(X_i^{\mathrm{T}}\beta)-Y_i\exp(-X_i^{\mathrm{T}}\beta)\}=0 \tag{6}$$

记方程(6)的解为 $\hat{\beta}$. 下面,我们将研究 $\hat{\beta}$ 的渐近性质.

在研究渐近性质之前,我们需要定义一些记号. 对于向量或者矩阵 a，记 $a^{\otimes 2}=a\,a^{\mathrm{T}}$. 定义 $\psi_i(\beta)=X_i\{T_i^{-1}\exp(X_i^{\mathrm{T}}\beta)-T_i\exp(-X_i^{\mathrm{T}}\beta)\}$，定义 T_i 的生存函数为 $K(u)=P(T_i>u)$. 假设跟踪研究时间为 L. 由于计数过程和鞅理论在证明中被用到,我们需要相关的概念. 定义 σ 流 $\mathscr{F}(u)$ 由如下 σ 代数生成：

$$\sigma\{I(C_i\leqslant t),t\leqslant u;I(T_i\leqslant y),X_i,\\ 0\leqslant y\leqslant L,i=1,\cdots,n\}.$$

记 C_i 的计数过程为 $N_i^c(u)=I(Y_i\leqslant u,\delta_i=0)$，风险过程为 $R_i(u)=I(Y_i\geqslant u)$，删失分布的风险函数为 $\lambda^c(u)$. 在随机删失假设下，$\mathscr{M}_i^c(u)=N_i^c(u)-\int_0^u\lambda^c(t)R_i(t)\mathrm{d}t$ 是一个鞅. 定义 $\mathscr{M}^c(u)=\sum_{i=1}^{n}\mathscr{M}_i^c(u)$，$N^c(u)=\sum_{i=1}^{n}N_i^c(u)$ 和 $R(u)=\sum_{i=1}^{n}R_i(u)$.

为了保证相合性和渐近正态性,我们需要假设下列正则条件.

正则条件

(C1) 失效时间 T_i 被删失时间 C_i 随机删失,即 (T_i,X_i,ε_i) 与删失时间 C_i 是独立的.

(C2) 存在一个 $\delta_0>0$，使得 $E[(1/\varepsilon_i+\varepsilon_i)^2\exp(\delta_0\|X_i\|)]<\infty$.

(C3) $E(X_iX_i^{\mathrm{T}})$ 是正定的.

(C4) 误差项 ε_i 满足 $E(\varepsilon_i-1/\varepsilon_i|X_i)=0$.

(C5) $T_i\leqslant L$ 且 $P(C_i\geqslant L)>0$.

注　条件(C1)是逆概率加权方法所需要的条件,可参考文献[16-17] 等. 条件(C2)保证存在一个包含 β_0 的紧集 Θ 使得估计函数及其导数在 Θ 上是一致收敛的.条件(C3)是估计方程中经常用到的条件.条件(C4)是文献[4]中的可识别条件,为了保证所建立的估计方程是渐近无偏的,类似于最小二乘和最小一乘中的误差的期望为零和中位数为零.若 $\ln\varepsilon$ 的分布关于 0 对称,则条件(C4) 满足. 条件(C5)是估计平均寿命和剩余寿命常用的条件,可以参考文献[18].

定理 1.1　在条件(C1)～(C5)下,在 β_0 的邻域内存在估计 $\hat{\beta}$ 使得 $\hat{\beta}$ 以概率收敛到 β_0. 而且，

$$\sqrt{n}(\hat{\beta}-\beta_0)\xrightarrow{d}N(0,V_0^{-1}DV_0^{-1}),$$

其中，

$$V_0=E[X_iX_i^{\mathrm{T}}(\varepsilon_i+1/\varepsilon_i)],$$

$$H(\psi,u)=\frac{1}{K(u)}E[\psi_i(\beta_0)I(u\leqslant T_i)],$$

$$D=E[\psi_i(\beta_0)^{\otimes 2}]+E\left[\int_0^L\frac{\{H(\psi,u)-\psi_i(\beta_0)\}^{\otimes 2}}{S(u)^2}\lambda^c(u)R_i(u)\mathrm{d}u\right]:=D_1+D_2.$$

定理 1.1 的证明见节 2.

由于(X_i,ε_i)，C_i 的分布未知,参数 β 的真实值β_0 未知,因此,$\hat{\beta}$ 的方差需要用样本进行估计.V_0 和 D_1 可以用经验的样本均值进行估计，即

$$\hat{V}_0=\frac{1}{n}\sum_{i=1}^{n}\frac{\delta_i}{\hat{S}(Y_i)}X_iX_i^{\mathrm{T}}\{Y_i\exp(-X_i^{\mathrm{T}}\hat{\beta})+Y_i^{-1}\exp(X_i^{\mathrm{T}}\hat{\beta})\},$$

$$\hat{D}_1=\frac{1}{n}\sum_{i=1}^{n}\frac{\delta_i}{\hat{S}(Y_i)}X_iX_i^{\mathrm{T}}\{Y_i\exp(-X_i^{\mathrm{T}}\hat{\beta})-Y_i^{-1}\exp(X_i^{\mathrm{T}}\hat{\beta})\}^2.$$

由于 $\psi_i(\beta_0)$中含有不能完全观测的 T_i，我们可以得到

$$D_2=E\left[\int_0^L\frac{\{H(\psi,u)-\psi_i(\beta_0)\}^{\otimes 2}}{S(u)^2}\lambda^c(u)I(T_i\geqslant u)I(C_i\geqslant u)\mathrm{d}u\right]=\\ E\left[\int_0^L\frac{\{H(\psi^{\otimes 2},u)-H(\psi,u)^{\otimes 2}\}}{S(u)^2}\lambda^c(u)I(R_i\geqslant u)\mathrm{d}u\right],$$

其中，

$$H(\psi^{\otimes 2},u)=(1/K(u))E[\psi_i(\beta_0)^{\otimes 2}I(T_i\geqslant u)].$$

因此,D_2 可以用下式来估计,即

$$\hat{D}_2=n^{-1}\sum_{i=1}^{n}\int_0^L\frac{\mathrm{d}N_i^c(u)}{\hat{S}(u)^2}\{\hat{H}(\psi^{\otimes 2},u)-\hat{H}(\psi,u)^{\otimes 2}\}=$$

$$n^{-1}\sum_{i=1}^{n}\frac{1-\delta_i}{\hat{S}(Y_i)^2}\{\hat{H}(\psi^{\otimes 2},Y_i)-\hat{H}(\psi,Y_i)^{\otimes 2}\},$$

其中,

$$\hat{H}(\psi,u)=\frac{1}{n\hat{K}(u^-)}\sum_{i=1}^{n}\frac{\delta_i}{\hat{S}(Y_i)}\psi_i(\hat{\beta})I(Y_i\geqslant u),$$

$$\hat{H}(\psi^{\otimes 2},u)=\frac{1}{n\hat{K}(u^-)}\sum_{i=1}^{n}\frac{\delta_i}{\hat{S}(Y_i)}\psi_i(\hat{\beta})^{\otimes 2}I(Y_i\geqslant u).$$

从而 $V_0^{-1}DV_0^{-1}$ 可以通过 $\hat{V}_0^{-1}(\hat{D}_1+\hat{D}_2)\hat{V}_0^{-1}$ 来估计.

2 定理证明

定理 1.1 的证明 记

$$U_n(\beta)=\sum_{i=1}^{n}\frac{\delta_i}{\hat{S}(Y_i)}X_i\{Y_i^{-1}\exp(X_i^{\mathrm{T}}\beta)-Y_i\exp(-X_i^{\mathrm{T}}\beta)\},$$

$$L_n(\beta)=\sum_{i=1}^{n}\frac{\delta_i}{\hat{S}(Y_i)}\{Y_i^{-1}\exp(X_i^{\mathrm{T}}\beta)+Y_i\exp(-X_i^{\mathrm{T}}\beta)\},$$

$$\phi(\beta)=E\,[T_i^{-1}\exp(X_i^{\mathrm{T}}\beta)+T_i\exp(-X_i^{\mathrm{T}}\beta)].$$

第一步:相合性

通过一些简单的代数计算,我们可以得到 $U_n(\beta)=\partial L_n(\beta)/\partial\beta$.又 $L_n(\beta)$是严格凸的,因此,方程(6)的解就是 $L_n(\beta)$的最小值点. 又

$$\frac{1}{n}L_n(\beta)=\frac{1}{n}\sum_{i=1}^{n}\frac{\delta_i\{S(Y_i)-\hat{S}(Y_i)\}}{S(Y_i)\hat{S}(Y_i)}\cdot\{Y_i^{-1}\exp(X_i^{\mathrm{T}}\beta)+Y_i\exp(-X_i^{\mathrm{T}}\beta)\}+\frac{1}{n}\sum_{i=1}^{n}\frac{\delta_i}{S(Y_i)}\cdot\{Y_i^{-1}\exp(X_i^{\mathrm{T}}\beta)+Y_i\exp(-X_i^{\mathrm{T}}\beta)\}:=I_{1n}+I_{2n}.$$

又 $T_i^{-1}\exp(X_i^{\mathrm{T}}\beta)=\varepsilon_i^{-1}\exp\{X_i^{\mathrm{T}}(\beta-\beta_0)\}\leqslant\varepsilon_i^{-1}\exp(\|\beta-\beta_0\|\cdot\|X_i\|)$. 类似的,我们可以得到

$$T_i^{-1}\exp(X_i^{\mathrm{T}}\beta)+T_i\exp(-X_i^{\mathrm{T}}\beta)\leqslant(\varepsilon_i^{-1}+\varepsilon_i)\exp(\|\beta-\beta_0\|\cdot\|X_i\|)\tag{7}$$

由条件(C2)可得,我们可以找到一个包含 β_0 的紧集 Θ 使得

$$(\varepsilon_i^{-1}+\varepsilon_i)\exp(\|\beta-\beta_0\|\cdot\|X_i\|)\leqslant(\varepsilon_i^{-1}+\varepsilon_i)\exp(\delta_0\|X_i\|)\tag{8}$$

结合式(7),(8)和一致大数定理,可得

$$\sup_{\beta\in\Theta}\|T_i^{-1}\exp(X_i^{\mathrm{T}}\beta)+T_i\exp(-X_i^{\mathrm{T}}\beta)-\phi(\theta)\|\xrightarrow{p}0\tag{9}$$

记 $Y_{(n)}$ 为 $Y_1,\cdots,Y_n$ 中最大的次序统计量,由 Kaplan-Meier 估计的一致相合性[19]和式(9),可以得出

$$|I_{2n}|\leqslant\sup_{0\leqslant y\leqslant Y_{(n)}}\frac{|S(y)-\hat{S}(y)|}{S(y)\hat{S}(y)}\cdot\frac{1}{n}\sum_{i=1}^{n}\{T_i^{-1}\exp(X_i^{\mathrm{T}}\beta)+T_i\exp(-X_i^{\mathrm{T}}\beta)\}=o_p(1)\tag{10}$$

关于 $\beta\in\Theta$ 是一致的. 又经过一些计算,我们可以得到

$$E\left[\frac{\delta_i}{S(Y_i)}\{Y_i^{-1}\exp(X_i^{\mathrm{T}}\beta)+Y_i\exp(-X_i^{\mathrm{T}}\beta)\}\right]=E\left[E\left[\frac{\delta_i}{S(T_i)}\{T_i^{-1}\exp(X_i^{\mathrm{T}}\beta)+T_i\exp(-X_i^{\mathrm{T}}\beta)\}\mid T_i,X_i\right]\right]=E\left[\frac{1}{S(T_i)}\left\{T_i^{-1}\exp(X_i^{\mathrm{T}}\beta)+T_i\exp(-X_i^{\mathrm{T}}\beta)\right\}E\left[\delta_i\,\middle|\,T_i,X_i\right]\right]=\phi(\beta).$$

因此,类似(9)的证明,可得 $I_{2n}=\phi(\beta)+o_p(1)$,关于 $\beta\in\Theta$ 是一致的.

综合以上结果,可得 $L_n(\beta)/n=\phi(\beta)+o_p(1)$,关于 $\beta\in\Theta$ 是一致的.而且又因为

$$\partial^2\phi(\beta)/\partial\beta\,\partial\beta^{\mathrm{T}}=E\left[X_iX_i^{\mathrm{T}}\left\{T_i^{-1}\exp(X_i^{\mathrm{T}}\beta)+T_i\exp(-X_i^{\mathrm{T}}\beta)\right\}\right]\geqslant E[X_iX_i^{\mathrm{T}}]$$

是正定的,所以 $\phi(\beta)$是严格凸的.由条件(C4)可得,$\partial\phi(\beta)/\partial\beta\,|_{\beta=\beta_0}=0$,从而有 β_0 是 $\phi(\beta)$的最小值点. 由文献[20,定理 5.7]可得,$\hat{\beta}$ 依概率收敛到 β_0.

第二步:渐近正态性

$$U_n(\beta_0)=\sum_{i=1}^{n}\frac{\delta_i}{S(Y_i)}\psi_i(\beta_0)+\sum_{i=1}^{n}\frac{\delta_i\{S(Y_i)-\hat{S}(Y_i)\}}{S(Y_i)\hat{S}(Y_i)}\psi_i(\beta_0):=A_1+A_2.$$

运用下面的等式[21]:

$$\frac{\delta_i}{S(Y_i)}=1-\int_0^L\frac{\mathrm{d}\mathcal{M}_i^c(u)}{S(u)},$$

A_1 可以被分解为

$$A_1=\sum_{i=1}^{n}\psi_i(\beta_0)-\sum_{i=1}^{n}\int_0^t\psi_i(\beta_0)\frac{\mathrm{d}\mathcal{M}_i^c(u)}{S(u)}.$$

Gill[22]证明了

$$\frac{\widehat{S}(t)-S(t)}{S(t)}=-\int_0^t\frac{\widehat{S}(u^-)}{S(u)}\frac{\mathrm{d}\mathscr{M}^c(u)}{R(u)} \tag{11}$$

式中，$\widehat{S}(u^-)$是删失时间生存函数的 Kaplan-Meier 估计的左连续版本. 又注意到

$$n^{-1}R(u)=\widehat{K}(u^-)\widehat{S}(u^-) \tag{12}$$

式中，$\widehat{K}(u)$是 $K(u)=P(T>u)$的 Kaplan-Meier 估计.由式(11)和(12)，我们可以得到

$$\frac{S(t)-\widehat{S}(t)}{S(t)}=\int_0^t\frac{\widehat{S}(u^-)}{S(u)}\frac{\mathrm{d}\mathscr{M}^c(u)}{R(u)}=$$
$$\int_0^L\frac{\widehat{S}(u^-)I(u\leqslant t)}{S(u)}\frac{\mathrm{d}\mathscr{M}^c(u)}{R(u)}=$$
$$\int_0^L\frac{I(u\leqslant t)}{nS(u)\widehat{K}(u^-)}\mathrm{d}\mathscr{M}^c(u).$$

因此

$$A_2=\sum_{i=1}^n\frac{\delta_i\psi_i(\beta_0)}{\widehat{S}(Y_i)}\int_0^L\frac{I(u\leqslant Y_i)}{nS(u)\widehat{K}(u^-)}\mathrm{d}\mathscr{M}^c(u)=$$
$$\int_0^L\frac{1}{n\widehat{K}(u^-)}\sum_{i=1}^n\frac{\delta_i\psi_i(\beta_0)I(u\leqslant T_i)}{\widehat{S}(T_i)}\frac{\mathrm{d}\mathscr{M}^c(u)}{S(u)}.$$

由一致大数定律和 Kaplan-Meier 估计的一致相合性[19]可得

$$\sup_{0\leqslant u\leqslant L}\left\|\frac{1}{n\widehat{K}(u^-)}\sum_{i=1}^n\frac{\delta_i\psi_i(\beta_0)I(u\leqslant T_i)}{\widehat{S}(T_i)}-H(\psi,u)\right\|\xrightarrow{p}0.$$

从而可得

$$A_2=\sum_{i=1}^n\int_0^L\frac{H(\psi,u)}{S(u)}\mathrm{d}\mathscr{M}_i^c(u)+o_p(n^{1/2}).$$

综上所述，可以得到

$$U_n(\beta_0)=\sum_{i=1}^n\psi_i(\beta_0)+$$
$$\sum_{i=1}^n\int_0^L\frac{H(\psi,u)-\psi_i(\beta_0)}{S(u)}\mathrm{d}\mathscr{M}_i^c(u)+o_p(n^{1/2}) \tag{13}$$

由于 $\psi_i(\beta_0)$关于 $\mathscr{F}(0)$是可测的，因此，式(13)右侧的前两项是不相关的. 从而，由鞅中心极限定理可得

$$n^{-1/2}U_n(\beta_0)\xrightarrow{d}N(0,D).$$

从而，由 Taylor 展开可以得到

$$0=U_n(\widehat{\beta})=U_n(\beta_0)+V_n(\beta^*)n(\widehat{\beta}-\beta_0),$$

式中，$V_n(\beta)=\frac{1}{n}\partial U_n(\beta)/\partial\beta^{\mathrm{T}}$. 经过一些简单的运算，我们可以得到

$$V_n(\beta)=$$
$$\frac{1}{n}\sum_{i=1}^n\frac{\delta_i}{\widehat{S}(Y_i)}X_iX_i^{\mathrm{T}}\{Y_i^{-1}\exp(X_i^{\mathrm{T}}\beta)+Y_i\exp(-X_i^{\mathrm{T}}\beta)\}.$$

$V_n(\beta)$可以分解为

$$V_n(\beta)=\frac{1}{n}\sum_{i=1}^n\frac{\delta_i}{S(Y_i)}X_iX_i^{\mathrm{T}}\cdot$$
$$\{Y_i^{-1}\exp(X_i^{\mathrm{T}}\beta)+Y_i\exp(-X_i^{\mathrm{T}}\beta)\}+$$
$$\frac{1}{n}\sum_{i=1}^n\frac{\delta_i\{S(Y_i)-\widehat{S}(Y_i)\}}{S(Y_i)\widehat{S}(Y_i)}X_iX_i^{\mathrm{T}}\cdot$$
$$\{Y_i^{-1}\exp(X_i^{\mathrm{T}}\beta)+Y_i\exp(-X_i^{\mathrm{T}}\beta)\}:=$$
$$V_{1n}(\beta)+V_{2n}(\beta).$$

又对于任意常数 $C>0$，我们有

$$\|X_i\|^2=\|X_i\|^2I(\|X_i\|^2\leqslant C)+$$
$$C\frac{\|X_i\|^2}{C}I(\|X_i\|^2>C)\leqslant$$
$$C+C\exp(2\|X_i\|/\sqrt{C})\leqslant$$
$$2C\exp(2\|X_i\|/\sqrt{C}) \tag{14}$$

结合式(7)和(14)，存在一个足够大的常数 C 和 C_1，使得

$$\|X_iX_i^{\mathrm{T}}\{T_i^{-1}\exp(X_i^{\mathrm{T}}\beta)+T_i\exp(-X_i^{\mathrm{T}}\beta)\}\|\leqslant$$
$$2C\exp(2\|X_i\|/\sqrt{C})\cdot$$
$$(\varepsilon_i^{-1}+\varepsilon_i)\exp(\|\beta-\beta_0\|\cdot\|X_i\|)\leqslant$$
$$C_1(\varepsilon_i^{-1}+\varepsilon_i)\exp(\delta_0\|X_i\|) \tag{15}$$

对任意的 $\beta\in\Theta$ 成立.类似于(10)的证明，我们可以得到 $V_{2n}=o_p(1)$关于 $\beta\in\Theta$ 是一致的.令 $V(\beta)=E[X_iX_i^{\mathrm{T}}\{T_i^{-1}\exp(X_i^{\mathrm{T}}\beta)+T_i\exp(-X_i^{\mathrm{T}}\beta)\}]$. 经过简单计算，我们有

$$E\left[\frac{\delta_i}{S(Y_i)}X_iX_i^{\mathrm{T}}\{Y_i^{-1}\exp(X_i^{\mathrm{T}}\beta)+\right.$$
$$\left.Y_i\exp(-X_i^{\mathrm{T}}\beta)\}\right]=V(\beta).$$

由一致大数定律可以得到

$$\sup_{\beta\in\Theta}\|V_n(\beta)-V(\beta)\|\xrightarrow{p}0.$$

又 $V(\beta_0)=V_0$ 是正定的，$\widehat{\beta}$ 是相合的，因此，可以得到

$$\sqrt{n}(\widehat{\beta}-\beta_0)=$$
$$-V_n^{-1}(\beta^*)n^{-1/2}U_n(\beta_0)\xrightarrow{d}N(0,V_0^{-1}DV_0^{-1}).$$

此定理证明结束.

3　数值模拟

在本节中，我们进行一些数值模拟来评估我们所提方法的效果. 从下面的模型中随机得到生成响应变量 T_i，

$$T_i=\exp(c_0+\alpha_1X_{i1}+\alpha_2X_{i2})\varepsilon_i,\ i=1,\cdots,n,$$

式中，协变量 X_{i1} 和 X_{i2} 相互独立且服从均值为 0，

方差为 1 的正态分布，$(c_0, \alpha_1, \alpha_2)=(1,1,2)$. 我们考虑两个随机误差项 ε_i 的分布：$\ln\varepsilon_i \sim N(0,1)$ 和 $\ln\varepsilon_i \sim U(-2,2)$. 删失时间 C_i 由一个取值为无穷的单点分布和一个均值为 10 的指数分布的混合分布生成. 通过调整这两个分布的混合比例，我们可以得到大约 20％～25％的删失比例. 在模拟中，我们将考虑 3 种方法：Naive 方法，加权最小二乘方法(WLS) 和我们提出的方法(Proposed). 其中，Naive 方法是对完全观测的样本使用普通最小二乘方法；WLS 方法是指 Stute[12-13] 所提出的方法；Proposed 方法是指节 1 所提出的方法.

为了评估 Naive 方法，WLS 方法和 Proposed 的方法，我们计算了偏差(Bias)、标准差(SE)、标准差的估计(SEE)和置信水平为 95％的置信区间的覆盖率(CP). 其中，Naive 方法的 SEE 是通过Plug-in 方法计算得到，WLS 和 Proposed 方法的 SEE 是基于 1000 次的 bootstrap 得到的. 样本量 n 取 100、200，我们重复 1000 次模拟，模拟结果由表 1 给出.

从表 1 可以看出，Naive 方法对参数的估计偏差很大；WLS 方法和 Proposed 方法有较小的偏差，且偏差和标准差随着样本的增大而减小. 当 $\ln\varepsilon_i \sim N(0,1)$时，WLS 方法是有效的，Proposed 方法的表现效果跟 WLS 方法类似. 当 $\ln\varepsilon_i \sim U(-2,2)$时，Proposed 方法有较小的方差，比 WLS 方法表现好. 另外，标准差的估计(SEE)和标准差(SE)很接近，95％置信区间的覆盖率接近 95％. 因此，从模拟中可以看出我们提出的方法在有限样本下有较好的表现.

表 1 20％～25％删失比例下的模拟结果

Tab.1 Simulation results under the 20％～25％ censoring rate

n	method		$\ln\varepsilon \sim U(-2,2)$			$\ln\varepsilon \sim N(0,1)$		
			$c_0(1)$	$\alpha_1(1)$	$\alpha_1(2)$	$c_0(1)$	$\alpha_1(1)$	$\alpha_1(2)$
100	Naive	Bias	−0.167	−0.037	−0.076	−0.128	−0.024	−0.057
		SE	0.137	0.137	0.147	0.120	0.116	0.123
		SEE	0.136	0.135	0.140	0.117	0.116	0.121
		CP	0.757	0.928	0.895	0.793	0.946	0.925
	WLS	Bias	−0.006	−0.003	0.005	−0.009	0.000	0.003
		SE	0.136	0.142	0.145	0.117	0.124	0.133
		SEE	0.137	0.143	0.145	0.120	0.124	0.127
		CP	0.941	0.943	0.935	0.952	0.949	0.933
	Proposed	Bias	−0.006	−0.004	0.003	−0.013	−0.002	0.001
		SE	0.115	0.120	0.122	0.125	0.130	0.137
		SEE	0.117	0.123	0.125	0.120	0.122	0.124
		CP	0.947	0.955	0.940	0.929	0.933	0.922
200	Naive	Bias	−0.175	−0.039	−0.072	−0.123	−0.031	−0.058
		SE	0.093	0.096	0.100	0.085	0.082	0.088
		SEE	0.095	0.094	0.097	0.083	0.082	0.085
		CP	0.562	0.922	0.871	0.671	0.932	0.879
	WLS	Bias	−0.006	−0.004	0.001	−0.001	0.002	0.002
		SE	0.098	0.103	0.104	0.082	0.089	0.091
		SEE	0.096	0.100	0.102	0.085	0.087	0.089
		CP	0.942	0.933	0.929	0.957	0.932	0.943
	Proposed	Bias	−0.007	−0.003	−0.001	−0.006	0.002	0.000
		SE	0.082	0.087	0.087	0.090	0.094	0.096
		SEE	0.080	0.084	0.086	0.087	0.088	0.089
		CP	0.935	0.940	0.938	0.941	0.928	0.924

参考文献(References)

[1] CHEN K N, GUO S J, LIN Y Y, et al. Least absolute relative error estimation [J]. Journal of the American Statistical Association, 2010, 105(491): 1104-1112.

[2] ZHANG Q Z, WANG Q H. Local least absolute relative error estimating approach for partially linear multiplicative model [J]. Statistica Sinica, 2013, 23: 1091-1116.

[3] ZHANG T, ZHANG Q Z, LI N X. Least absolute relative error estimation for functional quadratic multiplicative model [J]. Communications in Statistics: Theory and Methods, 2016, 45(19): 5802-5817.

[4] CHEN K N, LIN Y Y, WANG Z F, et al. Least product relative error estimation [J]. Journal of Multivariate Analysis, 2016, 144: 91-98.

[5] WANG Z F, LIU W X, LIN Y Y. A change-point problem in relative error-based regression [J]. Test, 2015, 24: 835-856.

[6] LIU X H, LIN Y Y, WANG Z F. Group variable selection for relative error regression [J]. Journal of Statistical Planning & Inference, 2016, 175: 40-50.

[7] MILLER R G. Least squares regression with censored data [J]. Biometrika, 1976, 63: 449-464.

[8] BUCKLEY J, JAMES I. Linear regression with censored data[J]. Biometrika, 1979, 66(3): 429-436.

[9] JAMES I R, SMITH P J. Consistency results for linear regression with censored data [J]. The Annals of Statistics, 1984, 12(2): 590-600.

[10] JIN Z Z, LIN D Y, YING Z L. On least-squares regression with censored data [J]. Biometrika, 2006, 93(1): 147-161.

[11] KOUL H, SUSARLA V, VAN RYZIN J. Regression analysis with randomly right-censored data [J]. The Annals of Statistics, 1981, 9(6): 1276-1288.

[12] STUTE W. Consistent estimation under random censorship when covariables are present [J]. Journal of Multivariate Analysis, 1993, 45: 89-103.

[13] STUTE W. Distributional convergence under random censorship when covariables are present [J]. Scandinavian Journal of Statistics, 1996, 23: 461-471.

[14] HE S Y, HUANG X. Central limit theorem of linear regression model under right censorship [J]. Science in China Series A: Mathematics, 2003, 46(5): 600-610.

[15] BAO Y C, HE S Y, MEI C L. The Koul-Susarla-Van Ryzin and weighted least squares estimates for censored linear regression model: A comparative study [J]. Computational Statistics & Data Analysis, 2007, 51: 6488-6497.

[16] BANG H, TSIATIS A A. Estimating medical costs with censored data [J]. Biometrika, 2000, 87(2): 329-343.

[17] MA Y Y, YIN G S. Censored quantile regression with covariate measurement errors [J]. Statistica Sinica, 2011, 21: 949-971.

[18] SUN L Q, SONG X Y, ZHANG Z G. Mean residual life models with time-dependent coefficients under right censoring [J]. Biometrika, 2012, 99(1): 185-197.

[19] FLEMING T R, HARRINGTON D P. Counting Processes and Survival Analysis [M]. New York: Wiley, 1991.

[20] VAN DER VAART A W. Asymptotic Statistics [M]. Cambridge: Cambridge University Press, 1998.

[21] ROBINS J M, ROTNITZKY A, ZHAO L P. Estimation of regression coefficients when some regressors are not always observed [J]. Journal of the American Statistical Association, 1994, 89(427): 846-866.

[22] GILL R D. Censoring and Stochastic Integrals [M]. Amsterdam: Mathematisch Centrum, 1980.