第 34 卷 第 5 期 华侨大学学报(自然科学版) Vol. 34 No. 5
2013 年 9 月 Journal of Huaqiao University (Natural Science) Sep. 2013

文章编号:1000-5013(2013)05-0591-05 doi:10.11830/ISSN.1000-5013.2013.05.0591

随机删失数据下核密度估计的相合性

叶彩园,吴群英,刘振

(桂林理工大学 理学院,广西 桂林 541006)

摘要: 在$\{X_i,i\geqslant 1\}$是α混合的随机变量,$\{Y_i,i\geqslant 1\}$是独立同分布的随机变量,且X_i与Y_i相互独立的情形下,研究随机删失数据下概率密度函数的核估计,获得此核估计的逐点强相合性和一致强相合性.

关键词: 删失数据;核密度估计;逐点强相合性;一致强相合

中图分类号: O 22.4 **文献标志码:** A

核密度估计方法最早由 Rosenblat 等[1]于 1956 年引入. 在生存分析中,由于实验设计、观察时间等因素的局限,所要收集的数据往往不能完全观测,这便产生了删失数据. 因此,研究随机删失数据的大样本性质是解决实际问题的关键. 对于独立非删失数据样本,$\hat{f}_n(x)$的相合性已取得了较深入的研究. 例如在独立样本下,熊丹[2]讨论了核密度估计的r阶均方相合性及收敛速度;陈希儒等[3]讨论了$\hat{f}_n(x)$的相合性. 对于大多数的金融、时间序列而言,样本的独立性不一定成立;而样本的相依性是它们固有的特征,混合结构,尤其是α-混合结构能比较合理的刻画时间序列模型的相依结构. 对于混合非删失数据样本,国内外的许多专家对$\hat{f}_n(x)$的相合性进行了讨论. 例如李军[4]在φ混合样本下,讨论了$\hat{f}_n(x)$的强相合性;韦来生[5]在 NA 样本下讨论了 的相合性;赵翌等[6]在α混合序列下讨论了核密度估计量的强相合性与一致强相合性. 对于混合删失数据样本,王海建等[7]研究了被删失变量和删失变量之间不独立的情况下,$f_n(x)$的估计形式及密度函数估计的强一致收敛速度;Liang 等[8]等研究了强混合删失样本下核密度估计的 Berry-Esseen 界;Cai[9-10]研究了删失相依数据 Kaplan-Meier 估计的渐进正态性,以及核密度函数和失效率函数的估计问题;Wang[11]研究了基于截尾数据概率密度核估计的一些渐进行为;张成毅等[12]讨论了缺失数据下局部估计的弱项合性和渐进正态性;石小平[13]利用经验生存函数和子经验函数,研究了删失变量的生存函数估计的相合性,并讨论了其收敛性. 然而,对于在删失样本与被删失样本相互独立的情形下,$f_n(x)$的相合性问题还未见文献报道. 本文在α-混合序列下,研究了随机删失数据下概率密度函数的核估计,获得了此核估计的逐点强相合性和一致强相合性.

1 相关定义和引理

令$\{X_n,n\geqslant 1\}$为概率空间(Ω,F,P)上实值随机变量序列,$F_n^m=\sigma(X_i;n\leqslant i\leqslant m)$是由随机变量序列$\{X_i;n\leqslant i\leqslant m\}$产生的$\sigma$代数,定义

$$\alpha(n)=\sup_{k\geqslant 1}\sup_{A\in F_1^k,B\in F_{k+n}^{\infty}}|P(AB)-p(A)P(B)|.$$

若$\lim\limits_{n\to\infty}\alpha(n)=0$,则称随机变量序列$\{X_n;n\geqslant 1\}$为$\alpha$混合. 文中出现的$c$均表示与$n$无关的常数.

引理 1[8] 假设$EX_i=0$且$|X_i|\leqslant S<\infty$, a. s. $(i=1,2,\cdots,n)$,对于$n,m\in N$,$0<m<n/2$,$\varepsilon>0$,有

$$P(|\sum_{i-1}^{n}X_i|<\varepsilon)\leqslant 4\exp\{-\frac{\varepsilon^2}{16}(nm^{-1}D_m+\frac{1}{3}\varepsilon Sm)^{-1}\}+32\frac{S}{\varepsilon}n\alpha(m),$$

收稿日期: 2013-03-26

通信作者: 吴群英(1961-),女,教授,主要从事概率统计的研究. E-mail:wqy666@glut.edu.cn.

基金项目: 国家自然科学基金资助项目(11061012);广西高校人才小高地建设创新团队计划项目(桂教人(2011)47号);广西研究生教育创新计划项目(YCSZ2012086)

其中：$D_m = \max\limits_{1\leqslant j\leqslant 2m} \mathrm{var}(\sum\limits_{i=1}^{j} X_i)$.

引理 2[8] 假设 $\alpha(k)\leqslant C_1 k^{-\gamma}$，$\gamma>1$，若 $\sup\limits_{1\leqslant j,j\leqslant n,i\neq j}|\mathrm{cov}(X_i,X_j)|\triangleq R^*(n)<\infty$，存在 m，$2\gamma/(\gamma-1)<m<\infty$，使得 $R_m(n)=\sup\limits_{1\leqslant i\leqslant n}(E|X_i|^m)^{1/m}<\infty$，且 $R_\infty(n)=\sup\limits_{1\leqslant i\leqslant n}\ \mathrm{ess}_{\omega\in\Omega}|X_i|$，则有

$$\mathrm{var}(\sum_{i=1}^{n} X_i)\leqslant n\{C_2(\gamma,m)(R_m(n))^{2m/(\gamma(m-2))}(R^*(n))^{1-m/(\gamma(m-2))}+R_2^2(n)\}.$$

其中：$C_2(\gamma,m)=\dfrac{20\gamma-40\gamma/m}{\gamma-1-2\gamma/m}C_1^{1/\gamma}$.

引理 3[3] 设 $K(\cdot)$及 $g(\cdot)$均为 R^1 上 Borel 可测函数，满足条件：1) K 有界；2) $\int|K(u)|\mathrm{d}u<\infty$；3) $\lim\limits_{|u|\to\infty} uK(u)=0$；4) $\int|g(u)|\mathrm{d}u<\infty$. 常数序列$\{h_n\}$满足$\lim\limits_{n\to\infty}h_n=0$. 令 $g_n(x)=\dfrac{1}{h_n}\times\int K(u/h_n)g(x-u)\mathrm{d}u$，则对任意 $x\in C(g)$，$C(g)$为 g 的连续点集，都有$\lim\limits_{n\to\infty}g_n(x)=g(x)\int K(u)\mathrm{d}u$. 进一步，若 g 有界且一致连续，则上式关于 x 一致成立.

引理 4[9] 如果$\{X_n;n\geqslant1\}$是一列具有混合系数 $\alpha(n)$的随机变量，与独立同分布的随机变量列$\{Y_n;n\geqslant1\}$相互独立，则$\{(X_n,Y_n);n\geqslant1\}$是一列具有混合系数为 $4\alpha(n)$的随机变量. 特别地，$\{(X_n\wedge Y_n);n\geqslant1\}$也是混合系数为 $4\alpha(n)$的随机变量.

2 主要结果

设 $X_1,X_2,\cdots,X_n$ 是具有同分布的 α-混合随机变量，具有连续的分布函数 F 和相应的密度函数 f. $Y_1,Y_2,\cdots,Y_n$ 是独立同分布表示删失的随机变量，具有连续分布函数 G，且 X_i 与 Y_i 独立.

为了证明定理，给出以下几个基本假设.

K1) $\{X_n;n\geqslant1\}$的混合系数 $\alpha(n)$满足 $\alpha(n)=O(n^{-v})$，$v>3$.

K2) 任给整数 $j\geqslant1$，X_j 在给定 $X_i=x_1$ 条件下的条件密度记为 $f_{j-i}(\cdot|x_1)$，且设存在常数 $M_0>0$，使对于 $x\in R$，存在 $x_1,x_2\in U(x)$，有 $f_j(x_1|x_2)\leqslant M_0$.

K3) 若 K 在 R^1 有界且满足：ⅰ) $\int_{-\infty}^{\infty}K(u)\mathrm{d}u=1$；ⅱ) $\int_{-\infty}^{\infty}uK(u)\mathrm{d}u=1$；ⅲ) $\lim\limits_{u\to\infty}uK(u)=0$；ⅳ) $\int_{-\infty}^{\infty}u^2K(u)\mathrm{d}<\infty$.

定理 1 设 $X_1,X_2,\cdots,X_n$ 是具有同分布的 α-混合随机变量，连续的密度函数 $f(x)$；$Y_1,Y_2,\cdots,Y_n$ 是独立同分布表的随机变量，且 X_i 与 Y_i 相互独立. 若基本假设 K1)～K3)成立，存在 β，满足 $\beta v>1$，且存在 $r_n\to\infty$，使得

$$\lim_{n\to\infty}\frac{r_n(\ln n)^{\beta+1}}{n^{1-1/v}h_n^{1+1/v}}=0, \tag{1}$$

则对 $x<\tau_H$，有$|f_n(x)-f(x)|\to0$，a.s..

定理 2 设 $X_1,X_2,\cdots,X_n$ 是具有同分布的 α-混合随机变量，$Y_1,Y_2,\cdots,Y_n$ 是独立同分布的随机变量，且 X_i 与 Y_i 独立. 若基本假设 K1)～K3)成立，且 $f(x)$和 $K(\cdot)$都满足 Lipschitz 条件，当取 $h_n=n^{-l}$，$0<l<\min(\dfrac{v-1-\eta}{v+3},\dfrac{v-1}{3})$，$0<\eta<v-1-3l$ 时，对于 $\tau<\tau_H$，有 $\sup\limits_{x\in(0,\tau)}|f_n(x)-f(x)|\to0$，a.s..

3 定理证明

3.1 定理 1 的证明

显然有

$$|f_n(x)-f(x)|\leqslant|Ef_n(x)-f(x)|+|f_n(x)-Ef_n(x)|\triangleq I_1+I_2.$$

欲证$|f_n(x)-f(x)|\to0$，a.s.，只需证明$|Ef_n(x)-f(x)|\xrightarrow{\text{a.s.}}0$ 和$|f_n(x)-Ef_n(x)|\xrightarrow{\text{a.s.}}0$ 即可.

利用条件 K3)中的 i)，$Z_1,Z_2,\cdots,Z_n$ 的同分布性及引理 3，可得

$$I_1=|Ef_n(x)-f(x)|=\left|\frac{1}{nh_n}\sum_{i=1}^{n}EK(\frac{Z_i-x}{h_n})\frac{\delta_i}{1-G(Z_i)}-f(x)\right|=$$

$$\left|\frac{1}{h_n}\int_{-\infty}^{\infty}K(\frac{s-x}{h_n})f(s)\mathrm{d}s-f(x)\right|=\left|\frac{1}{h_n}\int_{-\infty}^{\infty}K(\frac{u}{h_n})[f(x+u)-f(x)]\mathrm{d}u\right|\to 0. \tag{2}$$

$$I_2=|f_n(x)-Ef_n(x)|=$$

$$\left|\sum_{i=1}^{n}\frac{1}{nh_n}[K(\frac{Z_i-x}{h_n})\frac{\delta_i}{1-G(Z_i)}-EK(\frac{Z_i-x}{h_n})\frac{\delta_i}{1-G(Z_i)}]\right|\triangleq\left|\sum_{i=1}^{n}\eta_i\right|.$$

其中：$\eta_i=\frac{1}{nh_n}\{K(\frac{Z_i-x}{h_n})\frac{\delta_i}{1-G(Z_i)}-EK(\frac{Z_i-x}{h_n})\frac{\delta_i}{1-G(Z_i)}\}$.

根据引理 4，可知$\{\eta_i,i\geqslant1\}$也是 α-混合的，满足引理 2 的条件．故根据 C_r 不等式、基本条件 K2)，以及 $f(x)$，$G(x)$的连续性、X_i 与 Y_i 相互独立和引理 3 可知，对于 $x<\tau_H$，$i\neq j$，有

$$|\mathrm{cov}(\eta_i,\eta_j)|=|E\eta_i\eta_j|\leqslant\frac{c}{(nh_n)^2}\left|E[K(\frac{Z_i-x}{h_n})\frac{\delta_i}{1-G(Z_i)}K(\frac{Z_j-x}{h_n})\frac{\delta_j}{1-G(Z_j)}]\right|\leqslant$$

$$\frac{c}{(nh_n)^2}\left|E[K(\frac{X_i-x}{h_n})\frac{1}{1-G(X_i)}K(\frac{X_j-x}{h_n})\frac{1}{1-G(X_j)}]\right|\leqslant$$

$$\frac{c}{(nh_n)^2}\left|\iint K(\frac{s-x}{h_n})\frac{1}{1-G(s)}K(\frac{t-x}{h_n})\frac{1}{1-G(t)}f(s)f_{j-i}(t\mid s)\mathrm{d}s\mathrm{d}t\right|=$$

$$\frac{c}{n^2}\left|\int_{-\infty}^{\infty}K(u)f(x+h_nu)\mathrm{d}u\right|=O(n^{-2}). \tag{3}$$

根据 C_r 不等式、$f(x)$，$G(x)$的连续性，以及 K 的有界性和引理 3，可得

$$R_m(n)=\sup_{1\leqslant i\leqslant n}\left(\left|\frac{1}{nh_n}[K(\frac{Z_i-x}{h_n})\frac{\delta_i}{1-G(Z_i)}-EK(\frac{Z_i-x}{h_n})\frac{\delta_i}{1-G(Z_i)}]\right|^m\right)^{1/m}\leqslant$$

$$\left\{\frac{c}{(n^mh_n^{m-1})^m}\left|\int_{-\infty}^{\infty}K(u)f(x+h_nu)\mathrm{d}u\right|\right\}^{1/m}=O(n^{-1}h_n^{1/m-1}), \tag{4}$$

$$R_\infty(n)=\sup_{1\leqslant i\leqslant n}\operatorname*{ess}_{\omega\in\Omega}|\eta_i|\leqslant\frac{1}{nh_n}. \tag{5}$$

根据引理 2 及式(3)～(5)，可得

$$D_m=\max_{1\leqslant j\leqslant 2m}\mathrm{var}(\sum_{i=1}^{j}\eta_i)\leqslant cm[n^{-1}h_n^{1/m-1})^{2m/r(m-2)}n^{-2(1-m/r(m-2)}+n^{-2}h_n^{-1}]=O(mn^{-2}h_n^{-1}). \tag{6}$$

取 $m=h_n^{-1/v}n^{1/v}(\ln n)^\beta$，根据引理 1 及式(1)，(6)，由于有 $\beta v>1$，故可得

$$P(|f_n(x)-Ef_n(x)|>\varepsilon)=P(|\sum_{i-1}^{n}\eta_i|>\varepsilon)\leqslant$$

$$\exp\{-cn^{-1}h_n^{-1}+\frac{(\ln n)^\beta}{n^{1-1/v}h_n^{1+1/v}})^{-1}\}+\frac{1}{n(\ln n)^{\beta v}}\leqslant$$

$$\exp(-cr_n\ln n)+\frac{1}{n(\ln n)^{\beta v}}\leqslant\frac{1}{n^2}+\frac{1}{n(\ln n)^{\beta v}}.$$

则有

$$\sum_{n=1}^{\infty}P(|f_n(x)-Ef_n(x)|>\varepsilon)\leqslant\sum_{n=1}^{\infty}(\frac{1}{n^2}+\frac{1}{n(\ln n)^{\beta v}})<\infty.$$

由 Borel-Cantelli 引理可知

$$|f_n(x)-Ef_n(x)|\to 0,\qquad \text{a. s.}. \tag{7}$$

由式(2)，(7)可知，定理 1 得证．

3.2　定理 2 的证明

记 $I=[0,\tau]$，$\tau<\tau_H$，欲证 $\sup_{x\in(0,\tau)}|f_n(x)-f(x)|\to 0$，a. s.，显然有

$$\sup_{x\in I}|f_n(x)-f(x)|\leqslant\sup_{x\in I}|Ef_n(x)-f(x)|+\sup_{x\in I}|f_n(x)-Ef_n(x)|.$$

由于 $f(x)$在 R^1 满足 Lipschitz 条件，则 $f(x)$在 R^1 一致连续，根据条件 K3)中的 i) 和 $Z_1,Z_2,\cdots,$

Z_n 的同分布性及引理3可知

$$\sup_{x\in I}|Ef_n(x)-f(x)|=\sup_{x\in I}\left|\frac{1}{nh_n}\sum_{i=1}^{n}EK(\frac{Z_i-x}{h_n})\frac{\delta_i}{1-G(Z_i)}-f(x)\right|=$$
$$\sup_{x\in I}\left|\frac{1}{h_n}\int_{-\infty}^{\infty}K(\frac{s-x}{h_n})f(s)\mathrm{d}s-f(x)\right|=$$
$$\sup_{x\in I}\left|\frac{1}{nh_n}\int_{-\infty}^{\infty}K(\frac{u}{h_n})[f(x+u)-f(x)]\mathrm{d}u\right|\to 0. \tag{8}$$

将 $I=[0,\tau]$ 分成 N 个子区间,设 $0=x_1<x_2<\cdots<x_N=\tau$,第 i 个区间 I_i 记为 $\delta_i=\Delta x_i-x_{i-1}$,$i=1,2,\cdots,N$,$N\leqslant c\delta_n^{-1}$,且 $\delta_n=\max\limits_{1\leqslant i\leqslant N}\delta_i$. 取 δ_n 满足条件

$$\lim_{n\to\infty}\delta_n=0,\qquad \lim_{n\to\infty}\delta_n h_n^{-2}=0. \tag{9}$$

取 x_j^* 为 I_j 中的任一点,则有

$$\sup_{x\in I}|f_n(x)-Ef_n(x)|\leqslant\max_{1\leqslant j\leqslant N}\sup_{x\in I_j}|f_n(x)-f_n(x_j^*)|+$$
$$\max_{1\leqslant j\leqslant N}\sup_{x\in I_j}|Ef_n(x_j^*)-Ef_n(x)|+\max_{1\leqslant j\leqslant N}|f_n(x_j^*)-Ef_n(x_j^*)|\triangleq$$
$$I_{n,1}+I_{n,2}+I_{n,3}.$$

根据 $K(\cdot)$满足Lipschitz条件、式(9)及 $G(x)$的连续性可知

$$I_{n,1}=\max_{1\leqslant j\leqslant N}\sup_{x\in I_j}|f_n(x)-f_n(x_j^*)|=$$
$$\frac{1}{nh_n}\max_{1\leqslant j\leqslant N}\sup_{x\in I_j}\left|\sum_{i=1}^{n}[K(\frac{Z_i-x}{h_n})-K(\frac{Z_i-x_j^*}{h_n})]\frac{\delta_i}{1-G(Z_i)}\right|\leqslant$$
$$\frac{c}{nh_n}\max_{1\leqslant j\leqslant N}\sup_{x\in I_j}\sum_{i=1}^{n}\left|K(\frac{Z_i-x}{h_n})-K(\frac{Z_i-x_j^*}{h_n})\right|\leqslant$$
$$\frac{c}{h_n}\max_{1\leqslant j\leqslant N}\sup_{x\in I_j}\frac{|x_j^*-x|}{h_n}\leqslant c\delta_n h_n^{-2}\to 0. \tag{10}$$

由于 X_i 与 Y_i 相互独立且 $f(x)$满足Lipschitz条件,根据 $f(x)$,$G(x)$的连续性及式(9),可得

$$I_{n,2}=\max_{1\leqslant j\leqslant N}\sup_{x\in I_j}|Ef_n(x_j^*)-Ef_n(x)|=$$
$$\max_{1\leqslant j\leqslant N}\sup_{x\in I_j}\left|\frac{1}{h_n}\int_{-\infty}^{\infty}K(u)f(x_j^*+h_nu)\mathrm{d}u-\int_{-\infty}^{\infty}K(u)f(x+h_nu)\mathrm{d}u\right|\leqslant$$
$$\max_{1\leqslant j\leqslant N}\sup_{x\in I_j}\int_{-\infty}^{\infty}|K(u)|f(x_j^*+h_nu)-f(x-h_nu)|\mathrm{d}u\leqslant c\delta_n\to 0. \tag{11}$$

根据 $f_n(x)$的定义有

$$I_{n,3}=\max_{1\leqslant j\leqslant N}|f(x_j^*)-Ef(x_j^*)|=$$
$$\max_{1\leqslant j\leqslant N}\left|\sum_{i=1}^{n}\frac{1}{nh_n}[K(\frac{Z_i-x_j^*}{h_n})\frac{\delta_i}{1-G(Z_i)}-EK(\frac{Z_i-x}{h_n})\frac{\delta_i}{1-G(Z_i)}]\right|\triangleq$$
$$\max_{1\leqslant j\leqslant N}\left|\sum_{i=1}^{n}\xi_{i,j}\right|.$$

取 $m=n^{I+3l+\eta/v}(\ln n)^{\beta}$,根据引理3和式(6),由于有 $\beta v>1$,故可得

$$P(\max_{1\leqslant j\leqslant N}|f_n(x_j^*)-Ef_n(x_j^*)|>\varepsilon)=P(\max_{1\leqslant j\leqslant N}|\sum_{i=1}^{n}\xi_{i,j}|>\varepsilon)\leqslant$$
$$c\delta_n^{-1}\exp\{-\frac{\varepsilon^2}{16}(nm^{-1}nm^{-2}h_n^{-1}+\frac{1}{3}\varepsilon n^{-1}h_n^{-1}m)^{-1}\}+32\frac{n\alpha(m)}{\delta_n nh_n\varepsilon}\triangleq I_3+I_4.$$

由条件可知 $h_n=n^{-l}$,所以当取 $\delta_n=n^{-2l-\eta}$时,显然式(9)满足. 由于有 $0<l<\min(\frac{v-1-\eta}{v+3},\frac{v-1}{3})$,$0<\eta<v-1-3l$,则有 $1-l-\frac{1+3l+\eta}{v}>0$,所以有

$$I_3=c\delta_n^{-1}\exp\{-\frac{\varepsilon^2}{16}(nm^{-1}nm^{-2}h_n^{-1}+\frac{1}{3}\varepsilon n^{-1}h_n^{-1}m)^{-1}\}\leqslant$$

$$c\delta_n^{-1}\exp(-c(\ln n)^{\beta}n^{1-l-1+3l+\eta/v}) \leqslant$$

$$\frac{c}{n^2}cn^{-2l-\eta}\exp\{-(2+2l+\eta)\ln n\} \leqslant \frac{c}{n^2}, \tag{12}$$

$$I_4 = 32\frac{n\alpha(m)}{\delta_n nh_n\varepsilon} \leqslant c\frac{n^{-(1+3l+\eta)}(\ln\ \ n)^{\beta v}}{n^{-l}n^{-2l-\eta}} \leqslant \frac{c}{n(\ln\ \ n)^{\beta v}}. \tag{13}$$

则有

$$\sum_{n=1}^{\infty}P(\max_{1\leqslant j\leqslant N}|f_n(x_j^*)-Ef_n(x_j^*)|>\varepsilon) = \sum_{n=1}^{\infty}(\frac{c}{n^2}+\frac{c}{n(\ln n)^{\beta v}})<\infty.$$

故由 Borel-Cantelli 引理可知

$$\max_{1\leqslant j\leqslant N}|f_n(x_j^*)-Ef_n(x_j^*)|\to 0, \qquad \text{a. s.}. \tag{14}$$

根据式(8),(10),(11)和式(14)可知,定理 2 得证.

参考文献:

[1] ROSENBLAT M. Remarks on some nonparametric estimates of a density function[J]. Ann Math Statist,1956,27(3):832-837.

[2] 熊丹.独立样本核密度估计的 阶均方相合性及收敛速度[J].数学杂志,2004,24(3):303-306.

[3] 陈希孺,方兆本,李国英,陶波.非参数统计[M].上海:上海科学技术出版社,1989:263-264.

[4] 李军.φ-混合序列密度估计的强相合性[J].广西师范大学学报:自然科学版,1999,17(3):51-55.

[5] 韦来生.NA 样本概率密度函数核估计的相合性[J].系统科学与数学,2001,21(1):79-87.

[6] 赵翌,杨善朝.α 混合序列下的核密度估计量的强相合性[D].桂林:广西师范大学,2008:1-15.

[7] 王海建、赵跃生.删失数据密度函数估计及其强一致收敛速度[J].应用概率统计,1999,15(3):240-244.

[8] LIANG Han-ying,DE UNA-ALVAREZ J. A Berry-Essen type bound in kernel density estimation for strong mixing censored samples[J]. Journal of Multivariate Analysis,2009,100(6):1219-1231.

[9] CAI Zong-wu. Asymptotic properties of Kaplan-Meier estimator for censored dependent data[J]. Statistics & Probability Letters,1998,37(4):381-389.

[10] CAI Zong-wu. Kerner density and hazard rate estimation for censored dependent data[J]. Journal of Multivariate Analysis,1998,67(1):23-34.

[11] WANG Qi-hua. Some asymptotic behavior of the kernel estimates of a density function based on censored data[J]. Chin J Appl Probab Statist,1994,10(2):163-173.

[12] 张成毅,罗双华.缺失数据下局部估计的弱项合性和渐进正态性[J].纺织高校基础科学学报,2012,25(1):9-11.

[13] 石小平.删失变量生存函数的相合性及收敛性研究[D].乌鲁木齐:新疆大学,2006:15-22.

Consistency of Kernel Density Estimation for Randomly Censored Data

YE Cai-yuan, WU Qun-ying, LIU Zhen

(College of Science, Guilin University of Technology, Guilin 541006, China)

Abstract: In this paper, we study the estimation of a density function based on censored data by the kernel smoothing mothed when the survival and the censoring times form a stationary strong mixing sequence. Pointwise consistency and uniformly strong consistency of the kernel density estimation are derived.

Keywords: censored date; kernel density estimation; pointwise consistency; uniformly strong consistency

（**责任编辑**：陈志贤　　**英文审校**：黄心中）