Adv. & Appl. in Stat. **7**(2) (2007), 157-174

ESTIMATION OF THE CHANGE-POINTS OF THE MEAN RESIDUAL LIFE FUNCTION

SARA DIALLO

Université du Havre
LMAH, Faculté des Sciences et Techniques
B. P. 540, 76058 Le Havre Cedex, France
e-mail: sara.diallo617@univ-lehavre.fr

Abstract

Estimators for location and size of a discontinuity or change-point in a smooth mean residual life model are proposed. The proposed estimators also apply to the detection of discontinuities in derivatives and therefore to the detection of change-points of slope and of higher order curvature. The proposed estimators are based on a comparison of left and right one-sided kernel smoothers. Weak convergence of a stochastic process in local differences to a Gaussian process is established for properly scaled versions of estimators for the location of a change-point. The continuous mapping theorem can then be invoked to obtain asymptotic distributions and corresponding rates of convergence for change-point estimators.

1. Introduction

Nonparametric methods are usually applied in order to obtain a smooth fit of a curve without having to specify a parametric class of function. Sometimes a generally smooth curve might contain an isolated discontinuity or change-point in the curve or in a (possibly higher order) derivative, and in many cases interest focuses on the occurrence of such change-points.

2000 Mathematics Subject Classification: Primary 62G05; Secondary 62F12, 62G30.

Keywords and phrases: mean residual life function, jump size, kernel estimation, weak convergence, smoothing, Hille's theorem, survival function.

Received July 17, 2006

The analysis of change-points describing sudden, localized changes typically occurring in economics, medicine and the physical sciences has recently found increasing interest. General smoothness assumptions, allowing for a large class of functions to be considered, seem to be more appropriate in a variety of applied problems than parametric modeling.

The problem of estimating the location of the change-point of a mean residual life function (MRLF) is considered here.

Although the change-point has received relatively little attention in literature, some estimates of the change-point have been studied by Mi [11] and Ebrahimi [4]. We will consider estimates of the change-point of the MRLF via a kernel estimate of the MRLF. That is, if X is a real-valued random variable with the distribution function F, survival function $\overline{F}(x) = 1 - F(x)$ and such that $E(X^{+}) < \infty$; for example, X might represent the time of advice. The mean residual life function (MRLF for short) or the remaining life expectancy at age x, $M(x)$ of X is defined by (see, e.g., Kotz and Shanbhag [10], Hall and Wellner [8] and Guess and Proschan [7])

$$M(x) = E(X - x \mid X > x) = \begin{cases} \dfrac{\int_x^{\infty} \overline{F}(y) dy}{\overline{F}(x)} & \text{if } \overline{F}(x) > 0 \\ 0 & \text{otherwise.} \end{cases}$$

Given a sample $X_1, \ldots, X_n$ from F, then Yang [14] proposed a natural nonparametric estimate of $M(x)$ is the *random function* $e_n(x)$ defined by

$$e_n(x) = \frac{\int_x^{\infty} \overline{F}_n(y) dy}{\overline{F}_n(x)} \mathbb{I}_{[X_{(n)} > x]}, \tag{1.1}$$

where $X_{(n)} = \max_{1 \leq i \leq n} X_i$ that is the average, less x, of the observations exceeding x; and $\overline{F}_n(x)$ is the *empirical survival function* defined by $\overline{F}_n(x) = 1 - F_n(x)$, where $F_n(x) = \frac{1}{n} \sum_1^n \mathbb{I}_{[X_i \leq x]}$ is the empirical distribution. In order to introduce a kernel-type estimator for $M(x)$, let us use by $K(\cdot)$ (a probability density on the real line). Its corresponding survival function will be denoted by $\mathbb{K}(t) = \int_t^\infty K(u)du$. Also, we will need a sequence of smoothing parameters (or bandwidths) $h = h_n > 0$. The expected value of the empirical MRL estimator was derived by Abdous and Berred [1],

$$\mathbb{E}(e_n(x)) = e(x)(1 - F^n(x)). \tag{1.2}$$

It follows that the bias of the empirical estimator is $-e(x)F^n(x)$ and hence $e_n(x)$ is asymptotically unbiased, with bias decaying exponentially to zero as $n \to \infty$. When $\mathbb{E}(X^2) < \infty$, Abdous and Berred [1] also provided the variance of $e_n(x)$:

$$\mathcal{V}ar(e_n(x)) = e^2(x)F^n \times (1 - F^n(x))$$

$$+ \mathcal{V}ar[X - x \mid X > x] \sum_{j=1}^{n} \frac{1}{j} B(n, j, \overline{F}(x)), \tag{1.3}$$

where $B(n, j, \overline{F}(x)) = \binom{n}{j} \overline{F}^j \times F^{n-j}(x)$. Therefore, $\mathcal{V}ar(e_n(x)) \to 0$ as $n \to \infty$ when $\mathbb{E}(X^2) < \infty$.

Throughout this paper, we suppose that the MRLF M is l times continuously differentiable for some $l \geq 0$, $M \in \mathcal{C}^l$ and a kernel smoother with a kernel function of the order k is chosen, that is, the kernel function with exactly $(k - 1)$ vanishing moments.

However, the focus of this paper is to use the nonparametric regression method of the change-point to estimate the change-point of the MRLF. Smooth approximation of the change-point model by a model which contains a point of most rapid change and the corresponding statistical inference was considered by Müller and Wang [12] in the context of hazard functions under random censoring. Let $\nu \geq 0$ be an integer and $k \geq 2$ be an even integer. Assume that a change-point exists for $M^{(\nu)}$ at τ, $0 < \tau < 1$, in the following sense: There exists a

$g \in C^{(k+\nu)}([0, 1])$ such that

$$M^{(\nu)}(t) = g^{(\nu)}(t) + \Delta_\nu I_{[\tau, 1]}(t), \quad \Delta_\nu > 0, \, 0 \le t \le 1. \tag{1.4}$$

The case $\Delta_\nu < 0$ can be treated analogously. Define $M_+^{(\nu)}(\tau) = \lim_{x \downarrow \tau} M^{(\nu)}(x)$, $M_-^{(\nu)}(\tau) = \lim_{x \uparrow \tau} M^{(\nu)}(x)$ and $M^{(\nu)}(\tau) = M_+^{(\nu)}(\tau)$ and observe that

$$\Delta_\nu = M_+^{(\nu)}(\tau) - M_-^{(\nu)}(\tau), \tag{1.5}$$

where Δ_ν is the jump size at the possible change-point τ of the νth derivative. The case $\Delta_\nu = 0$ corresponds to the nonexistence of a change-point at τ. The change-point MRL function model (1.4) differs from the model of a change-point in the sequence of random variable (e.g., Hinkley [9], Deshayes and Picard [3], Worstley [13] in which the first k observations among $X_1, \ldots, X_k$ are independent and identically distributed with a common cumulative distribution function (c.d.f.) F while $X_{k+1}, \ldots, X_n$ are independent and identically distributed with a common c.d.f. G. In (1.4), the change-point τ is an unknown point in the domain of the common p.d.f. of all observations. In terms of the MRL function, this change-point is the unknown time at which the MRL function jumps.

The main results of this paper concern weak convergence of estimators $\hat{\tau}$ of the location of the change-point τ. The paper is organized as follows: Section 2 presents a discussion of kernel estimators using kernel functions with one-sided support and their application to change-point estimation, which is based on maximizing the difference between one-sided kernel smoothers. Section 3 is devoted to the study of a functional limit theorem for a local deviation process. The functional mapping theorem is used to obtain the distribution limit for the estimated change-point. The proofs of Section 3 are given in Section 4.

2. Change-point Estimators

We consider $M(x)$ as a function on its own, we can convolve $e_n(x)$

with $K_h(\cdot) = K(\cdot/h)/h$ and obtain the following kernel estimators:

$$\hat{M}^{(\nu)}(x) = \frac{1}{h^{\nu+1}} \int_{-\infty}^{+\infty} K^{(\nu)}\left(\frac{x-u}{h}\right) e_n(u) du. \tag{2.1}$$

Here $h = h_n$ is a sequence of bandwidths which is required to satisfy

$$h \to 0, \;\; nh^{2\nu+1} \to +\infty \;\; \text{as} \;\; n \to \infty, \;\; \lim_{n\to\infty} \sup nh^{2(k+\nu)+1} < \infty, \tag{2.2}$$

$K^{(\nu)}$ is the kernel function, which is assumed to be the νth derivative of a function K with compact support $[-1, 1]$ and $e_n(x)$ is defined by (1.1).

Let $K_+^{(\nu)}$ and $K_-^{(\nu)}$ be one-sided kernel functions with support $(K_+^{(\nu)}) = [-1, 0]$ and support $(K_+^{(\nu)}) = [0, 1]$ and define *one-sided estimates* for the νth derivative $M^{(\nu)}(x)$:

$$\hat{M}_{\pm}^{(\nu)}(x) = \frac{1}{h^{\nu+1}} \int_{-\infty}^{+\infty} K_{\pm}^{(\nu)}\left(\frac{x-u}{h}\right) e_n(u) du. \tag{2.3}$$

The idea is to base inference for change-points on differences between right and left sided estimates:

$$\hat{\Delta}^{(\nu)}(x) = \hat{M}_+^{(\nu)}(x) - \hat{M}_-^{(\nu)}(x). \tag{2.4}$$

Intuitively, the location of the maximum of these differences will be a reasonable estimator for the location of the change-point. Let $\mathcal{Q} \subset \,]0, 1[$ be a closed interval such that $\tau \in \mathcal{Q}$. Define the estimators

$$\hat{\tau} = \inf\{\rho \in \mathcal{Q} : \hat{\Delta}^{(\nu)}(\rho) = \sup_{x\in\mathcal{Q}} \hat{\Delta}^{(\nu)}(x)\} \tag{2.5}$$

for the location of the change-point τ and

$$\hat{\Delta}^{(\nu)}(\hat{\tau}) = \hat{M}_+^{(\nu)}(\hat{\tau}) - \hat{M}_-^{(\nu)}(\hat{\tau}) \tag{2.6}$$

for the jump size in the νth derivative. Defining $\hat{\tau}$ as maximizer over $\mathcal{Q}$ instead of $[0, 1]$ serves the sole purpose of excluding change-points located arbitrarily close to the boundary.

Assume that for some integer $\mu \geq 0$,

$$K \in \mathcal{C}^{\mu+\nu}([-1, 1]) \cap \mathcal{H}_{0,k}([-1, 1]), \tag{2.7}$$

$$K^{(j)}(-1) = K^{(j)}(1) = 0, \quad 0 \leq j < \mu + \nu, \tag{2.8}$$

where k as before is an even integer $k \geq 2$, $\nu < k$ and

$$\mathcal{H}_{\nu,l}([a_1, a_2]) = \left\{ \begin{array}{l} f \in \mathcal{C}([a_1, a_2]) : \text{support}(f) = [a_1, a_2], \\ \\ \int f(x) x^j dx = \begin{cases} = (-1)\nu!, & \text{if } j = \nu, \\ = 0, & \text{if } 0 \leq j <, \ j \neq \nu, \\ \neq 0, & \text{if } j = l, \end{cases} \end{array} \right\}.$$

It then follows by integration by parts that

$$K^{(\nu)} \in \mathcal{C}^{\mu}([-1, 1]) \cap \mathcal{H}_{\nu, k+\nu}([-1, 1]), \tag{2.9}$$

$$K^{(\nu+j)}(-1) = K^{(\nu+j)}(1) = 0, \quad 0 \leq j < \mu. \tag{2.10}$$

According to (2.10), the kernel $K^{(\nu)}$ is $(\mu - 1)$ times differentiable on $\mathbb{R}$ and $K^{(\mu-1)}$ is absolutely continuous. Similarly, assume for kernels K_+ and K_-,

$$K_+ \in \mathcal{C}^{\nu+\mu}([-1, 0]) \cap \mathcal{H}_{0,k}([-1, 0]), \tag{2.11}$$

$$K_- \in \mathcal{C}^{\nu+\mu}([0, 1]) \cap \mathcal{H}_{0,k}([0, 1]), \tag{2.12}$$

$$K_+^{(j)}(-1) = K_+^{(j)}(0) = 0, \quad 0 \leq j < \nu + \mu,$$

$$K_-^{(j)}(1) = K_-^{(j)}(0) = 0, \quad 0 \leq j < \nu + \mu, \tag{2.13}$$

which again imply that

$$K_+^{(\nu)} \in \mathcal{C}^{\mu}([-1, 0]) \cap \mathcal{H}_{\nu, k+\nu}([-1, 0]),$$

$$K_+^{(j+\nu)}(-1) = K_+^{(j+\nu)}(0) = 0, \quad 0 \leq j < \mu, \tag{2.14}$$

$$K_-^{(\nu)} \in \mathcal{C}^{\mu}([0, 1]) \cap \mathcal{H}_{\nu, k+\nu}([0, 1]),$$

$$K_-^{(j+\nu)}(1) = K_-^{(j+\nu)}(0) = 0, \quad 0 \leq j < \mu. \tag{2.15}$$

Observe that K_+ (respectively, K_-) acts on the right half side (r.h.s.) (respectively, left half side (l.h.s.)) of t according to the convolution property in definition (2.3), so that application of these kernels corresponds to employing smoothing windows $[t, t+h]$ (respectively, $[t-h, t]$). Observe that it follows from (2.13) that if $K_-^{(\nu)}$ satisfies (2.15), then a *kernel* $K_+^{(\nu)}$ defined by

$$K_+^{(\nu)}(x) = (-1)^{\nu} K_-^{(\nu)}(-x) \tag{2.16}$$

satisfies (2.14). An additional assumption we make is

$$K_-^{(\nu+\mu)}(0) > 0, \ (\nu+\mu) \text{ is odd and } \mu \geq 1. \tag{2.17}$$

Similar conditions follow for $K_+^{(\nu+\mu)}$, assuming (2.16).

3. Weak Convergence of Local Deviation Processes and Asymptotic Distributions of Change-point Estimators

In this section, a functional limit theorem for a process operating on increments of one-sided function estimates near τ is derived. The functional mapping theorem is then applied to obtain the limit distributions for change-point estimators $\hat{\tau}$. A similar device was used by Eddy [5, 6] in the context of estimating the mode of a probability density.

Let

$$\hat{\delta}_\nu(y) = \hat{\Delta}^{(\nu)}(\tau + yh) = \hat{M}_+^{(\nu)}(\tau + yh) - \hat{M}_-^{(\nu)}(\tau + yh)$$

and define for some $0 < T < \infty, -T \leq z \leq T$, the sequence of stochastic processes

$$\zeta_n(z) = (nh^{2\nu+1})^{(\mu+\nu+1)/(2(\mu+\nu))}\left(\hat{\delta}_\nu\left(\frac{z}{(nh^{2\nu+1})^{1/(2(\mu+\nu))}}\right) - \hat{\delta}_\nu(0)\right). \tag{3.1}$$

The scaling is chosen in such a way that processes ζ_n converge weakly.

Observe that $\zeta_n \in \mathcal{C}([-T, T])$. The following functional limit theorem holds.

Theorem 3.1. *Assume that* (1.1), (2.2) *and* (2.8)-(2.17) *hold. Then*

$$\zeta_n \to \zeta \quad a.s. \ \ on \ \ \mathcal{C}([-T, T]), \tag{3.2}$$

where ζ *is a continuous Gaussian process with moment structure*

$$\mathbb{E}(\zeta(z)) = -\frac{\Delta_\nu z^{\mu+\nu+1} K_-^{(\mu+\nu)}(0)}{(\mu+\nu+1)!}, \tag{3.3}$$

$$\operatorname{cov}(\zeta(z_1), \zeta(z_2)) = 2z_1 z_2 \sigma^2 \int (K_-^{(\nu+1)}(v))^2 dv, \tag{3.4}$$

where $\sigma^2 = Var(e_n)$.

Since the Gaussian limit process ζ is determined by its first and second moments, according to (3.3) and (3.4), it can be equally written as

$$\zeta(z) = -\frac{\Delta_\nu z^{\mu+\nu+1} K_-^{\mu+\nu}(0)}{(\mu+\nu+1)!} + Yz, \tag{3.5}$$

where $Y \sim \mathcal{N}\left(0,\ 2\sigma^2 \int (K_-^{(\mu+\nu)}(v))^2\, dv\right)$.

The proof of Theorem 3.1 follows from a sequence of lemmas in Section 4.

Asymptotic distributions of estimated change-points (2.5) can now be obtained as a consequence of this functional limit theorem. Under (2.17), the limit process ζ of (3.5) is seen to have a unique maximum at

$$Z^* = \left[\frac{Y(\mu+\nu)!}{\Delta_\nu K_-^{(\mu+\nu)}(0)}\right]^{1/(\mu+\nu)}. \tag{3.6}$$

Let Z_n be the location of the maximum of ζ_n. By construction,

$$\hat{\tau} = \tau + \frac{Z_n h}{(nh^{2\nu+1})^{1/(2(\mu+\nu))}}. \tag{3.7}$$

Corollary 3.1. *Under the assumption of Theorem* 3.1,

$$(nh^{2\nu+1})^{1/2}\left(\frac{\hat{\tau}-\tau}{h}\right)^{\mu+\nu} \to_{\mathcal{D}} \mathcal{N}\left(0, 2\left(\frac{(\mu+\nu)!}{\Delta_\nu (K_-^{(\mu+\nu)}(0))}\right)^2 \sigma^2 \int (K_-^{(\mu+\nu)}(v))^2 dv\right). \tag{3.8}$$

Consider for instance the important cases $\mu = 1, \ \nu = 0, \ k = 2$. If the usual bandwidth choice $h = dn^{-1/5}$ is made and $d > 0$, then (3.8) becomes

$$(n^{3/5}(\hat{\tau} - \tau) \to_{\mathcal{D}} \mathcal{N}\left(0, \ 2d\sigma^2 \frac{\int (K_-^{(1)}(v))^2 dv}{\Delta_\nu (K_-^{(1)}(0))^2}\right).$$

Another application of the functional mapping theorem shows that $\zeta_n(Z_n) \to_{\mathcal{D}} \zeta(Z^*)$ and therefore

$$(nh^{2\nu+1})^{1/2}\left\{\frac{\zeta_n(Z_n)}{(nh^{2\nu+1})^{(\mu+\nu+1)/(2(\mu+\nu))}}\right\} \to_{\mathcal{P}} 0.$$

This implies $(nh^{2\nu+1})^{1/2}\{\hat{\Delta}^{(\nu)}(\hat{\tau}) - \hat{\Delta}^{(\nu)}(\tau)\} \to_{\mathcal{P}} 0$, where $\hat{\Delta}^{(\nu)}(\cdot)$ is defined in (2.4). According to Lemma 4.6,

$$(nh^{2\nu+1})^{1/2}\{\hat{\Delta}^{(\nu)}(\tau) - \Delta_\nu\} \to_{\mathcal{D}} \mathcal{N}\left(0, \ 2\sigma^2 \int (K_-^{(\nu)}(v))^2 \, dv\right),$$

and combining these results one obtains for the jump size estimator $\hat{\Delta}^{(\nu)}(\hat{\tau})$.

Corollary 3.2.

$$(nh^{2\nu+1})^{1/2}\{\hat{\Delta}^{(\nu)}(\hat{\tau}) - \Delta_\nu\} \to_{\mathcal{D}} \mathcal{N}\left(0, \ 2\sigma^2 \int (K_-^{(\nu)}(v))^2 \, dv\right). \tag{3.9}$$

4. Auxiliary Results and Proofs

The following sequence of lemmas leads to the proof of Theorem 3.1.

Lemma 4.1.

$$\mathbb{E}(\zeta_n(z)) = -\frac{\Delta_\nu z^{\mu+\nu+1} K_-^{(\mu+\nu)}(0)}{(\mu + \nu + 1)!} + o(1). \tag{4.1}$$

Proof. Observe that, following (1.2) and (2.3),

$$\hat{M}_\pm^{(\nu)}(x, \ h) = \frac{1}{h^{\nu+1}} \int_{-\infty}^{+\infty} K_\pm^{(\nu)}\left(\frac{x - u}{h}\right) e_n(u) du,$$

then

$$\mathbb{E}(\hat{M}_{\pm}^{(\nu)}(\tau + yh)) = \frac{1}{h^{\nu+1}} \int K_{\pm}^{(\nu)}\left(\frac{\tau + yh - u}{h}\right)\mathbb{E}(e_n(u))du$$

$$= \underbrace{\frac{1}{h^{\nu}} \int K_{\pm}^{(\nu)}(v) M(\tau + yh - vh)dv}_{I}$$

$$- \underbrace{\frac{1}{h^{\nu}} \int K_{\pm}^{(\nu)}(v) M(\tau + yh - vh) F^{n}(\tau + yh - vh)dv}_{II}$$

$$= \frac{1}{h^{\nu}} \int K_{\pm}^{(\nu)}(v) M(\tau + yh - vh)dv + O([h^{\nu}]^{-1}).$$

Therefore, defining

$$\delta_{\nu}(y) = \frac{1}{h^{\nu}} \int_{-1}^{1} (K_{+}^{(\nu)}(v) - K_{-}^{(\nu)}(v)) M(\tau + yh - vh)dv,$$

we obtain

$$\mathbb{E}(\hat{\delta}_{\nu}(y)) = \delta_{\nu}(y) + O([h^{\nu}]^{-1}). \tag{4.2}$$

Observing (1.4), (2.11), (2.12), (2.13), evenness of k and (2.16) and employing a Taylor expansion and mean values $\xi_{1n} = \tau + \xi_1(y - v)h$, $\xi_{2n} = \tau + \xi_2(y - v)h$, then

$$\delta_{\nu}(y) = \frac{1}{h^{\nu}} \int_{-1}^{1} (K_{+}^{(\nu)}(v) - K_{-}^{(\nu)}(v))[M(\tau + yh - vh)(1_{\{v>y\}} + 1_{\{v\le y\}})]dv$$

$$= \underbrace{\frac{1}{h^{\nu}} \int_{-1}^{1} (K_{+}^{(\nu)}(v) - K_{-}^{(\nu)}(v))\left(\sum_{j=0}^{k+\nu-1} \frac{(y - v)^{j}}{j!} h^{j} M_{+}^{(j)}(\tau)1_{\{v\le y\}}\right)dv}_{(A)}$$

$$+ \underbrace{\frac{1}{h^{\nu}} \int_{-1}^{1} (K_{+}^{(\nu)}(v) - K_{-}^{(\nu)}(v))\left(\sum_{j=0}^{k+\nu-1} \frac{(y - v)^{j}}{j!} h^{j} M_{-}^{(j)}(\tau)1_{\{v>y\}}\right)dv}_{(B)}$$

$$+\underbrace{\frac{1}{h^{\nu}}\int_{-1}^{1}(K_{+}^{(\nu)}(v)-K_{-}^{(\nu)}(v))\frac{(y-v)^{k+\nu}}{(k+\nu)!}h^{k+\nu}M_{+}^{(k+\nu)}(\xi_{1n})1_{\{v\le y\}}dv}_{(C)}$$

$$+\underbrace{\frac{1}{h^{\nu}}\int_{-1}^{1}(K_{+}^{(\nu)}(v)-K_{-}^{(\nu)}(v))\frac{(y-v)^{k+\nu}}{(k+\nu)!}h^{k+\nu}M_{-}^{(k+\nu)}(\xi_{2n})1_{\{v>y\}}dv}_{(D)}$$

$$C+D=\frac{1}{h^{\nu}}\int_{-1}^{1}(K_{+}^{(\nu)}(v)-K_{-}^{(\nu)}(v))\frac{(y-v)^{k+\nu}}{(k+\nu)!}h^{k+\nu}(M_{+}^{(k+\nu)}(\xi_{1n})$$

$$-M_{+}^{(k+\nu)}(\tau)+M_{+}^{(k+\nu)}(\tau))1_{\{v\le y\}}dv$$

$$+\frac{1}{h^{\nu}}\int_{-1}^{1}(K_{+}^{(\nu)}(v)-K_{-}^{(\nu)}(v))\frac{(y-v)^{k+\nu}}{(k+\nu)!}h^{k+\nu}(M_{-}^{(k+\nu)}(\xi_{2n})$$

$$-M_{-}^{(k+\nu)}(\tau)+M_{-}^{(k+\nu)}(\tau))1_{\{v>y\}}dv$$

$$=\frac{1}{h^{\nu}}\int_{-1}^{1}(K_{+}^{(\nu)}(v)-K_{-}^{(\nu)}(v))\frac{(y-v)^{k+\nu}}{(k+\nu)!}h^{k+\nu}(M_{+}^{(k+\nu)}(\xi_{1n})$$

$$-M_{+}^{(k+\nu)}(\tau))\times 1_{\{v\le y\}}dv$$

$$+\frac{1}{h^{\nu}}\int_{-1}^{1}(K_{+}^{(\nu)}(v)-K_{-}^{(\nu)}(v))\frac{(y-v)^{k+\nu}}{(k+\nu)!}h^{k+\nu}(M_{-}^{(k+\nu)}(\xi_{2n})$$

$$-M_{-}^{(k+\nu)}(\tau))\times 1_{\{v>y\}}dv$$

$$+\frac{1}{h^{\nu}}\int_{-1}^{1}(K_{+}^{(\nu)}(v)-K_{-}^{(\nu)}(v))\frac{(y-v)^{k+\nu}}{(k+\nu)!}h^{k+\nu}M_{+}^{(k+\nu)}(\tau)1_{\{v\le y\}}dv$$

$$+\frac{1}{h^{\nu}}\int_{-1}^{1}(K_{+}^{(\nu)}(v)-K_{-}^{(\nu)}(v))\frac{(y-v)^{k+\nu}}{(k+\nu)!}h^{k+\nu}M_{-}^{(k+\nu)}(\tau)1_{\{v>y\}}dv.$$

Let

$$Q_n(y) = \frac{1}{h^\nu}\int_{-1}^{1}(K_+^{(\nu)}(v) - K_-^{(\nu)}(v))\frac{(y-v)^{k+\nu}}{(k+\nu)!}h^{k+\nu}(M_+^{(k+\nu)}(\xi_{1n})$$

$$- M_+^{(k+\nu)}(\tau)) \times 1_{\{v \le y\}}dv$$

$$+ \frac{1}{h^\nu}\int_{-1}^{1}(K_+^{(\nu)}(v) - K_-^{(\nu)}(v))\frac{(y-v)^{k+\nu}}{(k+\nu)!}h^{k+\nu}(M_-^{(k+\nu)}(\xi_{2n})$$

$$- M_-^{(k+\nu)}(\tau)) \times 1_{\{v > y\}}dv$$

$$= \frac{h^k}{(k+\nu)!}\int_{-1}^{1}(K_+^{(\nu)}(v) - K_-^{(\nu)}(v))(y-v)^{k+\nu}(M_+^{(k+\nu)}(\xi_{1n})$$

$$- M_+^{(k+\nu)}(\tau) + M_+^{(k+\nu)}(\tau))1_{\{v \le y\}}dv$$

$$+ \frac{h^k}{(k+\nu)!}\int_{-1}^{1}(K_+^{(\nu)}(v) - K_-^{(\nu)}(v))(y-v)^{k+\nu}(M_+^{(k+\nu)}(\xi_{2n})$$

$$- M_+^{(k+\nu)}(\tau) + M_-^{(k+\nu)}(\tau))1_{\{v > y\}}dv.$$

Then

$$A + B + C + D = Q_n(y)$$

$$+ \underbrace{\frac{1}{h^\nu}\int_{-1}^{1}(K_+^{(\nu)}(v) - K_-^{(\nu)}(v))\left(\sum_{j=0}^{k+\nu}\frac{(y-v)^j}{j!}h^j M_+^{(j)}(\tau)1_{\{v \le y\}}\right)dv}_{(E)}$$

$$+ \underbrace{\frac{1}{h^\nu}\int_{-1}^{1}(K_+^{(\nu)}(v) - K_-^{(\nu)}(v))\left(\sum_{j=0}^{k+\nu}\frac{(y-v)^j}{j!}h^j M_-^{(j)}(\tau)1_{\{v > y\}}\right)dv}_{(F)},$$

and by (2.11), (2.12) and (2.13), we have

$$F = \underbrace{\frac{1}{h^\nu}\sum_{j=0}^{k+\nu}\frac{h^j}{j!}M_-^{(j)}(\tau)\int_{-1}^{1}(K_+^{(\nu)}(v) - K_-^{(\nu)}(v))(y-v)^j 1_{\{v > y\}}dv}_{=0},$$

and

$$E = \frac{1}{h^{\nu}} \sum_{j=0}^{k+\nu} \frac{h^j (M_+^{(j)}(\tau) - M_-^{(j)}(\tau))}{j!} \int_{-1}^{1} (K_+^{(\nu)}(v) - K_-^{(\nu)}(v))(y-v)^j 1_{\{v \le y\}} dv$$

$$= \frac{(M_+^{(\nu)}(\tau) - M_-^{(\nu)}(\tau))}{\nu!} \int_{-1}^{1} (K_+^{(\nu)}(v) - K_-^{(\nu)}(v))(y-v)^{\nu} 1_{\{v \le y\}} dv$$

$$+ \underbrace{\frac{1}{h^{\nu}} \sum_{j=0}^{\nu-1} \frac{h^j (M_+^{(j)}(\tau) - M_-^{(j)}(\tau))}{j!} \int_{-1}^{1} (K_+^{(\nu)}(v) - K_-^{(\nu)}(v))(y-v)^j 1_{\{v \le y\}} dv}_{=0}$$

$$+ \underbrace{\frac{1}{h^{\nu}} \sum_{j=\nu+1}^{k+\nu} \frac{h^j (M_+^{(j)}(\tau) - M_-^{(j)}(\tau))}{j!} \int_{-1}^{1} (K_+^{(\nu)}(v) - K_-^{(\nu)}(v))(y-v)^j 1_{\{v \le y\}} dv}_{=0},$$

hence

$$\delta_{\nu}(y) = Q_n(y) + \frac{(M_+^{(\nu)}(\tau) - M_-^{(\nu)}(\tau))}{\nu!} \int_{-1}^{1} (K_+^{(\nu)}(v) - K_-^{(\nu)}(v))(y-v)^{\nu} 1_{\{v \le y\}} dv$$

$$= \frac{1}{\nu!} \Delta_{\nu} \int_{-1}^{1} (K_+^{(\nu)}(v) - K_-^{(\nu)}(v))(y-v)^{\nu} 1_{\{v \le y\}} dv + Q_n(y),$$

where $Q_n(y)$ is defined above.

Observe that

$$R_n(y) = | Q_n(y) - Q_n(0) | = o(h^k y), \tag{4.3}$$

since, for instance, for the difference of the first term on the r.h.s. of Q_n for $y > 0$,

$$\int_{-1}^{y} (K_+^{(\nu)}(v) - K_-^{(\nu)}(v))(-v)^{k+\nu} (M_+^{(k+\nu)}(\tau + \xi_1(-v)h) - M_+^{(k+\nu)}(\tau)) dv$$

$$= \int_{-1}^{y} (K_+^{(\nu)}(v-y) - K_-^{(\nu)}(v-y))(y-v)^{k+\nu} (M_+^{(k+\nu)}(\xi_{1n}) - M_+^{(k+\nu)}(\tau)) dv$$

and analogous calculations for $y < 0$, and for the difference of the second term on the r.h.s. of Q_n yields (4.3). Observing, for $y \geq 0$, under (2.16),

$$\frac{1}{\nu!}\int_0^y K_-^{(\nu)}(v)(y-v)^\nu dv = \int_0^y\left[\sum_{i=0}^{\mu-1}\frac{v^i}{i!}K_-^{(\nu+i)}(0)+\frac{v^\mu}{\mu!}K_-^{(\nu+\mu)}(\xi)\right]\frac{(y-v)^\nu}{\nu!}dv$$

$$= \frac{y^{\mu+\nu+1}}{(\mu+\nu+1)!}(K_-^{(\nu+\mu)}(0)+O(y)), \quad \text{as } y \to 0$$

and analogously, for $y \leq 0$,

$$\frac{1}{\nu!}\int_0^y K_-^{(\nu)}(v)(y-v)^\nu dv = \frac{-y^{\mu+\nu-1}}{(\mu+\nu+1)!}(K_-^{(\nu+\mu)}(0)+O(y)), \quad \text{as } y \to 0,$$

one obtains, noting that $K_-^{(\nu+\mu)}(0) = (-1)^{\mu+\nu}K_+^{(\nu+\mu)}(0)$ and that $(\mu+\nu)$ is odd

$$\delta_\nu(y)-\delta_\nu(0) = \frac{-\Delta_\nu K_-^{(\nu+\mu)}(0)y^{\mu+\nu+1}}{(\mu+\nu+1)!}(1+O(y)+o(h^k y)), \quad \text{as } y \to 0. \quad (4.4)$$

The result follows.

Lemma 4.2.

$$\operatorname{cov}(\zeta_n(z_1), \zeta_n(z_2)) = 2z_1z_2\sigma^2\int(K_-^{(\nu+1)}(v))^2dv + O\left(\frac{1}{(nh^{2\nu+1})^{1/2(\nu+\mu)}}\right). \quad (4.5)$$

Proof. Abbreviate $\alpha = 2\nu+1$, $\beta = \frac{(\mu+\nu+1)}{(2(\mu+\nu))}$ and $\gamma = \frac{1}{(2(\mu+\nu))}$.

$$\zeta_n(z)-\mathbb{E}(\zeta_n(z)) = \frac{(nh^\alpha)^\beta}{h^{\nu+1}}\int\Bigg[\left(K_+^{(\nu)}\left(\frac{\tau+(zh)/(nh^\alpha)^\gamma-v}{h}\right)-K_+^{(\nu)}\left(\frac{\tau-v}{h}\right)\right)$$

$$-\left(K_-^{(\nu)}\left(\frac{\tau+(zh)/(nh^\alpha)^\gamma-v}{h}\right)-K_-^{(\nu)}\left(\frac{\tau-v}{h}\right)\right)\Bigg]$$

$$\times(e_n(v)-\mathbb{E}(e_n(v)))dv. \quad (4.6)$$

This implies

$$\operatorname{cov}(\zeta_n(z_1), \zeta_n(z_2)) = \frac{(nh^{\alpha})^{2\beta}}{h^{2\nu+2}}\sigma^2$$

$$\times\left[\int\left[K_{+}^{(\nu)}\left(\frac{\tau+(z_1h)/(nh^{\alpha})^{\gamma}-v}{h}\right)-K_{+}^{(\nu)}\left(\frac{\tau-v}{h}\right)\right]dv\right.$$

$$\times\int\left[K_{+}^{(\nu)}\left(\frac{\tau+(z_2h)/(nh^{\alpha})^{\gamma}-v}{h}\right)-K_{+}^{(\nu)}\left(\frac{\tau-v}{h}\right)\right]dv$$

$$-\int\left[K_{+}^{(\nu)}\left(\frac{\tau+(z_1h)/(nh^{\alpha})^{\gamma}-v}{h}\right)-K_{+}^{(\nu)}\left(\frac{\tau-v}{h}\right)\right]dv$$

$$\times\int\left[K_{-}^{(\nu)}\left(\frac{\tau+(z_2h)/(nh^{\alpha})^{\gamma}-v}{h}\right)-K_{-}^{(\nu)}\left(\frac{\tau-v}{h}\right)\right]dv$$

$$-\int\left[K_{+}^{(\nu)}\left(\frac{\tau+(z_2h)/(nh^{\alpha})^{\gamma}-v}{h}\right)-K_{+}^{(\nu)}\left(\frac{\tau-v}{h}\right)\right]dv$$

$$\times\int\left[K_{-}^{(\nu)}\left(\frac{\tau+(z_1h)/(nh^{\alpha})^{\gamma}-v}{h}\right)-K_{-}^{(\nu)}\left(\frac{\tau-v}{h}\right)\right]dv$$

$$+\int\left[K_{-}^{(\nu)}\left(\frac{\tau+(z_1h)/(nh^{\alpha})^{\gamma}-v}{h}\right)-K_{-}^{(\nu)}\left(\frac{\tau-v}{h}\right)\right]dv$$

$$\left.\times\int\left[K_{-}^{(\nu)}\left(\frac{\tau+(z_2h)/(nh^{\alpha})^{\gamma}-v}{h}\right)-K_{-}^{(\nu)}\left(\frac{\tau-v}{h}\right)\right]dv\right].$$

By the assumptions, observing the compactness of supports,

$$K_{\pm}^{(\nu)}\left(\frac{\tau+(zh)/(nh^{\alpha})^{\gamma}-v}{h}\right)-K_{\pm}^{(\nu)}\left(\frac{\tau-v}{h}\right)$$

$$=K_{\pm}^{(\nu+1)}\left(\frac{\tau-v}{h}\right)\frac{z}{(nh^{\alpha})^{2\gamma}}$$

$$+O\left(\frac{1}{(nh^{\alpha})^{2\gamma}}\right)1_{\left\{K_{\pm}^{(\nu+1)}\left(\frac{\tau-v}{h}\right)\neq 0\right\}\cup\left\{K_{\pm}^{(\nu)}\left(\frac{\tau+(zh)/(nh^{\alpha})^{\gamma}-v}{h}\right)\neq 0\right\}}. \tag{4.8}$$

Inserting this into (4.7) and observing

$$1_{\left\{K_+^{(\nu+1)}\left(\frac{\tau-v}{h}\right)\neq 0\right\}\cup\left\{K_+^{(\nu)}\left(\frac{\tau+(zh)/(nh^\alpha)^\gamma-v}{h}\right)\neq 0\right\}} = O(h), \tag{4.9}$$

all the $O(\cdot)$ terms combined result in a summary $O(\cdot)$ term of $O(1/(nh^\alpha)^\gamma)$. Observing $2\beta - 2\gamma = 1$ and combining

$$\int K_\pm^{(\nu+1)}\left(\frac{\tau-v}{h}\right)du\int K_\pm^{(\nu+1)}\left(\frac{\tau-v}{h}\right)du = \frac{h}{n}\int (K_\pm^{(\nu+1)}(v))^2du + O\left(\frac{1}{n^2}\right),$$

$$\int K_\pm^{(\nu+1)}\left(\frac{\tau-v}{h}\right)du\int K_\pm^{(\nu+1)}\left(\frac{\tau-v}{h}\right)du = \frac{h}{n}\int K_+^{(\nu+1)}(v)K_-^{(\nu+1)}(v)dv + O\left(\frac{1}{n^2}\right)$$

$$= O\left(\frac{1}{n^2}\right) \tag{4.10}$$

with (4.7), where the differences are substituted by the leading terms of (4.8), completes the proof.

Lemma 4.3. *For fixed* z, $z \in [-T, T]$,

$$\zeta_n(z) - \mathbb{E}(\zeta_n(z)) \to_{\mathcal{D}} \mathcal{N}\left(0,\ 2z^2\sigma^2\int(K_-^{(\nu+1)}(v))^2dv\right).$$

Lemma 4.4. *For fixed* $z_1, z_2, \ldots, z_l$, $z_i \in [-T, T]$,

$$(\zeta_n(z_1) - \mathbb{E}(\zeta_n(z_1)), \ldots, \zeta_n(z_l) - \mathbb{E}(\zeta_n(z_l))) \to_{\mathcal{D}} \mathcal{N}(0, A), \tag{4.11}$$

where $A = (a_{ij})_{1\le i,j\le l}$ *and* $a_{ij} = 2z_iz_j\sigma^2\int(K_-^{(\nu+1)}(v))^2dv.$

Lemma 4.5. *The sequence* $\overline{\zeta}_n(\cdot) = \zeta_n(\cdot) - \mathbb{E}(\zeta_n(\cdot))$ *is tight.*

Proof. We show that there exists a constant $c > 0$ such that

$$\mathbb{E}(\overline{\zeta}_n(z_1) - \overline{\zeta}_n(z_2))^2 \le c(z_1 - z_2)^2 \tag{4.12}$$

for n sufficiently large. According to Billingsley [2], the moment condition (4.12) implies tightness $\overline{\zeta}_n$. Using the same notation as in the proof of Lemma 4.2 and defining

$$A_\pm(z) = \left\{u \in [0, 1] : K_\pm^{(\nu)}\left(\frac{\tau + zb/(nh^\alpha)^\gamma - u}{h}\right) \neq 0\right\},$$

the Lipschitz continuity of $K^{(\nu)}$ implies

$$\mathbb{E}(\overline{\zeta}_n(z_1) - \overline{\zeta}_n(z_2))^2$$

$$\leq \frac{(nh^{\alpha})^{2\beta}}{h^{2\nu+2}} \sigma^2 \left[\int \frac{|z_1 - z_2|}{(nh^{\alpha})^{\gamma}} (1_{A_+(z_1) \cup A_+(z_2)} + 1_{A_-(z_1) \cup A_-(z_2)}) du\right]^2$$

$$\leq c|z_1 - z_2|^2$$

since $2\beta - 2\gamma = 1$ and

$$\left[\int (1_{A_+(z_1) \cup A_+(z_2)} + 1_{A_-(z_1) \cup A_-(z_2)}) du\right]^2 = O(h/n^2).$$

Proof. Proof of Theorem 3.1.

Weak convergence of the processes $\overline{\zeta}_n$ follows now from applying Lemmas 4.4 and 4.5. The moment structure of the limit process ζ is a consequence of Lemmas 4.1 and 4.2.

The following lemma is used in the proof of Corollary 3.2.

Lemma 4.6.

$$(nh^{2\nu+1})^{1/2}\{\hat{\Delta}^{(\nu)}(\tau) - \Delta_\nu\} \to_{\mathcal{D}} \mathcal{N}\left(0,\ 2\sigma^2 \int (K_-^{(\nu)}(v))^2\, dv\right).$$

Acknowledgements

The author is grateful to Professor A. Berred for many valuable comments. The author would also like to thank the editors and the referee for a very careful reading of the manuscript and useful suggestions.

References

[1] Belkacem Abdous and Alexandre Berred, Mean residual life estimation, J. Statist. Plann. Inference 37 (2004), 327-337.

[2] P. Billingsley, Convergence of Probability Measures, Wiley, New York, 1968.

[3] J. Deshayes and D. Picard, Lois asymptotiques des test et estimateurs de rupture dans un modele statistique classique, Ann. Inst. H. Poincare 20 (1984), 309-327.

[4] Nader Ebrahimi, On estimating change point in a mean residual life function, Sankhya: The Indian Journal of Statistic 53 (1991), 206-219.

[5] W. F. Eddy, Optimum kernel estimators of the mode, Ann. Statist. 8 (1980), 870-882.

[6] W. F. Eddy, The asymptotic distributions of kernel estimators of the mode, Ann. Stat. 59 (1982), 279-290.

[7] F. Guess and F. Proschan, Mean residual life, Handbook of Statistics, Vol. 7, P. R. Rao Krishnaiah and C. R. Rao, eds., Amsterdam, North-Holland, 1988, pp. 215-224.

[8] W. J. Hall and J. A. Wellner, Mean residual life, Statistics and Related Topics, M. Csörgő, D. A. Dawson, J. N. K. Rao and A. K. Md. E. Saleh, eds., Amsterdam, North-Holland, 1981, pp. 169-184.

[9] D. V. Hinkley, Inference about the change-point in a sequence of random variables, Biometrika 57 (1970), 1-17.

[10] S. Kotz and D. N. Shanbhag, Some new approaches to probability distributions, Adv. Appl. Prob. 12 (1980), 903-921.

[11] Jie Mi, Estimation related to mean residual life, J. Nonparametr. Stat. 4 (1994), 179-190.

[12] H. G. Müller and J. L. Wang, Nonparametric analysis of changes in hazard rates for censored survival data: An alternative to change-point models, Biometrika 77 (1990), 305-314.

[13] K. J. Worstley, Confidence regions and tests for a change-point in a sequence of exponential family of random variables, Biometrika 73 (1986), 91-104.

[14] G. L. Yang, Estimation of a biometric function, Stochastic Process. Appl. 6 (1978), 112-116.

■