

SMOOTHING SIGNALS FOR SEMIMARTINGALES

A. THAVANESWARAN*

*Department of Statistics and Actuarial Science, University of Waterloo, Waterloo, Ontario, Canada
N2L 3G1*

Received 9 September 1985

Final revision 26 August 1987

The kernel function and convolution-smoothing methods developed to estimate a probability density function and distribution are essentially a way of smoothing the empirical distribution function. This paper shows now one can generalize these methods to estimate signals for a semimartingale model. A convolution-smoothed estimate is used to obtain an absolutely continuous estimate for an absolutely continuous signal of a semimartingale model. This provides a method of obtaining a convolution-smoothed estimate of the cumulative hazard function in the censored case, an open problem proposed by Mack (*Bulletin of Informatics and Cybernetics* 21 (1984) 29-35). Asymptotic properties of the convolution-smoothed estimate are discussed in some detail.

convolution-smoothing * kernel functions * semimartingales * signals * smoothing

1. Introduction

Let X_1, X_2, \dots, X_n be identically distributed independent random variables with density f and a distribution F . The corresponding empirical cumulative distribution function (e.c.d.f.) is

$$\begin{aligned} F_n(x) &= \text{proportion of observations} \leq x \\ &= \frac{1}{n} \sum_{i=1}^n I(X_i \leq x) \end{aligned}$$

where $I(A)$ denotes the indicator of the event A . Rosenblatt (1956) suggested that one might estimate the density by

$$f_n(x) = \int \frac{1}{b_n} w\left(\frac{x-t}{b_n}\right) dF_n(t) \quad (1)$$

where w is a function with integral 1, called the kernel function, and b_n is a positive parameter (the window). For a review of existing theory of density estimation by kernel functions, see Bean and Tsokos (1980).

* Currently with the Department of Biostatistics at the University of North Carolina at Chapel Hill. Research partly supported by the Office of Naval Research under Contract Number N00014-83-K-0387.

A fundamental problem with dependent observations is estimation of certain functionals of a “signal” process. In a recent paper Ramlau-Hansen (1985) focused on the intensity estimation for a counting process by means of kernel functions. He took as an analogy the case of kernel estimates of the hazard function for a distribution F concentrated on $[0, \infty)$; the hazard function is

$$\alpha(t) = \frac{f(t)}{1 - F(t)} \quad \text{for } t > 0 \text{ with } F(t) < 1.$$

By introducing $N(t) = nF_n(t)$ and $Y(t) = n - N(t-)$, the number at risk at time t , and setting $J(t) = I[Y(t) > 0]$, he treated the kernel function estimate as

$$\hat{\alpha}(t) = \frac{1}{b} \int_0^\infty w\left(\frac{t-s}{b}\right) d\hat{\beta}(s) \tag{2}$$

where $\hat{\beta}(t)$, given by the Stieltjes integral

$$\int_0^t \frac{1}{Y(s)} J(s) dN(s),$$

is the nonparametric estimator for the cumulative hazard function $\beta(t) = \int_0^t \alpha(s) ds$ introduced by Nelson (1972) and generalized by Aalen (1978); and it is seen that (2) essentially is a way of smoothing the increments in $\hat{\beta}(\cdot)$.

Semimartingale methods first came into prominence in survival analysis, where the observations are often heavily and individually censored, and the theory for the i.i.d. case does not apply. By applying the multiplicative intensity model for counting processes and stochastic integrals, Aalen (1978) demonstrated how it is possible to model such situations and develop nonparametric estimators for certain cumulative intensities. Aalen’s model did not incorporate more general jump processes or continuous path processes. Questions involving these led to the study of inference for a semimartingale model which includes diffusion processes, counting processes, gamma processes, Markov processes, semi-Markov processes etc.

In Section 2.1 we generalize to a semimartingale model of the form

$$dN(t) = \alpha(t) Y(t) dR_t + dM_t$$

where M_t is a locally square integrable martingale, and R_t is a continuous process of bounded variation. Some special cases are indicated, and the kernel estimator for $\alpha(t)$ is introduced.

The cumulative process $\beta^*(t)$ is $\int_0^t \alpha(s) J(s) dR_s$, where $J(t) = I[Y(t) > 0]$. Since $\beta^*(t)$ is absolutely continuous with respect to R_t , it is natural to look for an absolutely continuous estimator for $\beta^*(t)$. In Section 2.3 a convolution smoothed estimate for $\beta^*(t)$ is defined for a semimartingale model. Asymptotic properties of the convolution-smoothed estimate such as asymptotic normality and mean square uniform consistency are stated.

As a particular case we are concerned with the problem of estimating the cumulative hazard $\beta(t) = \int_0^t \alpha(s) ds$ for a counting process by applying the idea of Mack’s

(1984) convolution-smoothed estimator for the cumulative distribution function in the i.i.d. case when F is absolutely continuous. The question whether a convolution-smoothed estimate for censored/dependent observations can be defined remained open. A solution is given in Section 2.6 of this paper by applying the theory of previous sections to the appropriate counting process.

2. Smoothed estimates for semimartingales

2.1. Kernel estimates

In the interest of generality, consider the semimartingale model of the form

$$N(t) = \int_0^t \alpha(s) Y(s) dR_s + M_t, \quad \text{for } t \in [0, T]$$

where R_t is predictable and of bounded variation; M_t is a zero mean locally square integrable martingale with variance process $\langle M \rangle_t$, given by

$$\langle M \rangle_t = \int_0^t a(s) dR_s;$$

$a(t)$ and $Y(t)$ are non-negative predictable processes; $\alpha(t)$ is a deterministic unknown function (parameter); and the cumulative intensity $\beta^*(t) = \int_0^t J(s)\alpha(s) dR_s$ with $J(s) = I[Y(s) < 0]$ is assumed to be defined. Ramlau-Hansen (1983) considered the above model with M_t a martingale and $R_t = t$. It is of interest to note that

(i) the diffusion process model for which the kernel estimation is studied in Hasminskii and Ibragimov (1981) is a special case of the above model with $M_t = W_t$, the Wiener process, and $R_t = t$;

(ii) when M is a local martingale (martingale) the above model corresponds to Gill's (1980) (Aalen's (1978)) model;

(iii) when α , R are deterministic and M_t is an appropriately defined locally square integrable martingale then the above model corresponds to the extended gamma process model defined in Dykstra and Laud (1981), with $\alpha(t)$ and R_t representing the "scale" and "shape" parameters respectively.

Definition 2.1. Let w be a continuous function having compact support and integral 1, and let b be a positive parameter. The corresponding kernel estimator for the intensity α is given by the Stieltjes integral

$$\hat{\alpha}(t) = \frac{1}{b} \int_0^T w\left(\frac{t-s}{b}\right) d\hat{\beta}(s)$$

where $\hat{\beta}(t) = \int_0^t (J(s) dN(s)) / Y(s)$, the nonparametric estimate of $\beta^*(t)$.

2.2. Mean and variance of the kernel estimator

Let

$$\alpha^*(t) = \frac{1}{b} \int_0^T w\left(\frac{t-s}{b}\right) d\beta^*(s).$$

The mean of $\hat{\alpha}$, its variance, and an unbiased estimator of the variance are contained in the following proposition.

Proposition 2.2. *Assume that*

$$E \int_0^T w^2\left(\frac{t-s}{b}\right) \frac{J(s)}{Y^2(s)} d\langle M \rangle_s < \infty.$$

Then we have

$$\begin{aligned} E\hat{\alpha}(t) &= E\alpha^*(t), \\ \sigma^2(t) &= E[\hat{\alpha}(t) - \alpha^*(t)]^2, \\ &= \frac{1}{b^2} E \int_0^T w^2\left(\frac{t-s}{b}\right) \frac{J(s)}{Y^2(s)} d\langle M \rangle_s \end{aligned}$$

and an unbiased estimate of $\sigma^2(t)$ is given by

$$\hat{\sigma}^2(t) = \frac{1}{b^2} \int_0^T w^2\left(\frac{t-s}{b}\right) \frac{J(s)}{Y^2(s)} \cdot \frac{a(s)}{\alpha(s)Y(s)} dN(s).$$

Proof. Follows from the properties of stochastic integrals with respect to martingales. For a counting process with continuous compensator,

$$\langle M \rangle_t = \int_0^t \alpha(s)Y(s) ds \quad \text{and} \quad a(s) = \alpha(s)Y(s).$$

For the diffusion process model of the form (Hasminskii and Ibragimov (1981))

$$dX(t) = S(t) dt + \varepsilon dW(t) \quad \text{for } 0 \leq t \leq T,$$

the Cencov estimate $\hat{S}(t) = \int_0^T w(t-s) dX(s)$ corresponds to a kernel estimate with the Fejer kernel

$$w(u) = \frac{1}{2k} \left[\frac{\text{Sin}(\pi ku)}{\text{Sin}(\pi u)} \right]^2 \quad \text{where } k \rightarrow \infty \text{ as } \varepsilon \rightarrow 0.$$

It is easy to see that this is a special case of $\hat{\alpha}(t)$ above.

2.3. Convolution smoothed estimates

Here again we consider a semimartingale model of the form

$$N(t) = M(t) + \int_0^t \alpha(s)Y(s) dR_s,$$

as in Section 2.1, where $\beta^*(t) = \int_0^t \alpha(s)J(s) dR_s$, is the cumulative process to be estimated for $t \in [0, T]$. Following Mack (1984) a convolution-smoothed estimate can be motivated heuristically by defining $\tilde{\beta}(t) = \int_0^t \hat{\alpha}(u) dR_u$ where $\hat{\alpha}(u)$ is the kernel estimate of $\alpha(u)$ as in Section 1, given by

$$\hat{\alpha}(t) = \int_0^T \frac{1}{b} w\left(\frac{t-s}{b}\right) d\hat{\beta}(s)$$

with $\hat{\beta}(t)$ given by

$$\hat{\beta}(t) = \int_0^t \frac{J(s)}{Y(s)} dN(s).$$

Hence the convolution-smoothed estimate would be

$$\begin{aligned} \tilde{\beta}(t) &= \int_0^t \hat{\alpha}(u) dR_u = \frac{1}{b} \int_0^t \int_0^T w\left(\frac{u-s}{b}\right) \frac{J(s)}{Y(s)} dN(s) dR_u \\ &= \frac{1}{b} \int_0^T \left[\int_s^t w\left(\frac{u-s}{b}\right) dR_u \right] \frac{J(s)}{Y(s)} dN(s) = \int_0^T W\left(\frac{t-s}{b}\right) \frac{J(s)}{Y(s)} dN(s), \end{aligned}$$

which exists for example when

$$E \int_0^T W^2\left(\frac{t-s}{b}\right) \frac{J(s)}{Y^2(s)} d\langle M \rangle_s < \infty,$$

where

$$W\left(\frac{t-s}{b}\right) = \frac{1}{b} \int_0^t w\left(\frac{u-s}{b}\right) dR_u.$$

2.4. Properties of the convolution-smooth estimate

Let

$$\beta^{**}(t) = \int_0^T W\left(\frac{t-u}{b}\right) J(u)\alpha(u) dR_u.$$

Then we have the following proposition.

Proposition 2.3. *Let*

$$E \int_0^T W^2\left(\frac{t-u}{b}\right) \frac{J(u)}{Y^2(u)} d\langle M \rangle_u < \infty.$$

Then

- (i) $E[\tilde{\beta}(t)] = E[\beta^{**}(t)],$
- (ii)

$$(a)\sigma^2 = E[\tilde{\beta}(t) - \beta^{**}(t)]^2 = E \int_0^T W^2\left(\frac{t-u}{b}\right) \frac{J(u)}{Y(u)^2} d\langle M \rangle_u.$$

(b) An unbiased estimate of $\sigma^2(t)$ for a counting process model is

$$\hat{\sigma}^2(t) = \int_0^T W^2\left(\frac{t-u}{b}\right) \frac{J(u)}{Y^2(u)} dN(u).$$

In general (i) implies that the convolution-smoothed estimator is not an unbiased estimator of $\beta^*(t)$.

Proof

(i)

$$E[\tilde{\beta}(t) - \beta^{**}(t)] = E \int_0^T W\left(\frac{t-u}{b}\right) \frac{J(u)}{Y(u)} dM(u) = 0.$$

Hence the result.

(ii)

$$(a) \sigma^2(t) = E[\tilde{\beta}(t) - \beta^{**}(t)]^2 = E \left(\int_0^T W\left(\frac{t-u}{b}\right) \frac{J(u)}{Y(u)} dM(u) \right)^2.$$

Using a property of the Ito stochastic integral w.r.t. a martingale the R.H.S. equals

$$E \int_0^T W^2\left(\frac{t-u}{b}\right) \left[\frac{J(u)}{Y^2(u)} \right] d\langle M \rangle_u.$$

Hence the result.

The proof of (ii)b is similar to that of (i); note that when the model describes a counting process with continuous compensator $\langle M \rangle_t = \int_0^t \alpha(s) Y(s) ds$.

Suppose we consider a sequence $\{N_n\}$ of one-dimensional semimartingales each with the signal process of the form $\int_0^t \alpha(u) Y_n(u) dR_u$. We may for example think of N_n as the relevant counting process when the study population consists of n individuals or items. Let $J_n(s) = I[Y_n(s) > 0]$.

Proposition 2.4. Let $n \rightarrow \infty$ and b_n be such that $b_n \rightarrow 0$ as $n \rightarrow \infty$. Further if the intensity α is continuous at point t and if $j_n = EJ_n(t) \rightarrow 1$ uniformly in a neighbourhood of t , then

$$E\tilde{\beta}_n(t) \rightarrow E\beta^*(t). \quad (4)$$

Relation (4) follows from the definition of $\beta^{**}(t)$, and by (i) of Proposition 2.3 and by the fact that the sequence of functions $\{W((t-u)/b_n)\}$ is a Heaviside sequence as $n \rightarrow \infty$. i.e. Under the conditions stated, the convolution-smoothed estimator is asymptotically unbiased.

2.5. Asymptotic properties

Consider a sequence of semimartingales (N_n) on $[0, T]$ with a corresponding sequence of martingales given by

$$M_n(t) = N_n(t) - \int_0^t \alpha(s) Y_n(s) dR_s.$$

where $(\int_0^t (\alpha(s) Y_n(s)) dR_s)$ is the sequence of signal processes. The proofs of the following two theorems are similar to those of Ramlau-Hansen (1983), and hence are omitted. The first concerns asymptotic normality of the convolution smoothed estimator.

Theorem 2.5. *Assume that*

(i) $(nJ_n(s)/Y_n(s)) \xrightarrow{P} 1/\sigma(s)$

uniformly in a neighbourhood of t as $n \rightarrow \infty$,

(ii) *the functions α and σ are continuous at the point t .*

*Then $n^{1/2}[\tilde{\beta}_n(t) - \beta_n^{**}]$ converges in distribution to a normal distribution with mean 0 and variance*

$$\int_0^t \Delta(t-s) \frac{1}{\sigma(s)} dQ_s = \int_0^t \frac{1}{\sigma(s)} dQ_s \quad \text{where } Q_s = \lim_n \frac{\langle M_n \rangle_s}{Y_n(s)} \text{ in probability.}$$

Note $\Delta(t-s)$, a Heaviside function, is the limit of $W((t-s)/b_n)$ as $b_n \rightarrow 0$.

The next theorem gives mean square uniform consistency for the convolution-smoothed estimate.

Theorem 2.6. *Assume that*

(i) $J_n \rightarrow 1$ *in probability, uniformly on $[0, T]$ when $n \rightarrow \infty$,*

(ii) α *is continuous on $[0, T]$,*

(iii) $n\eta_n(T) = n \int_0^T E \left[\frac{J_n(s) d\langle M_n \rangle_s}{Y_n(s)} \right]$

is bounded when $n \rightarrow \infty$, and

(iv) W *is of bounded variation.*

Then

$$E \left[\sup_{t \in [a,b]} |\tilde{\beta}_n(t) - \beta(t)|^2 \right] \rightarrow 0 \text{ as } n \rightarrow \infty, \text{ } b_n, \text{ where } 0 < a < b < T.$$

2.6. Example

Consider independent identically distributed death or failure times X_1, \dots, X_n with values in $[0, \infty)$ and a hazard rate α , where

$$F(t) = 1 - \exp \left[- \int_0^t \alpha(s) ds \right] < 1.$$

Let T_1, \dots, T_n be corresponding independent identically distributed censoring times with distribution G . Assume that the censoring times are independent of the failure times. The number of failures

$$N_n(t) = \sum_{i=1}^n I(X_i \leq t, X_i \leq T_i)$$

is then a counting process with intensity process $\alpha(t)Y_n(t)$, where

$$Y_n(t) = \sum_{i=1}^n I(X_i \geq t, T_i \geq t)$$

denotes the number of individuals alive just before time t . The estimate for the cumulative hazard

$$\beta_n^*(t) = \int_0^t \alpha(s) I[Y_n(s) > 0] ds$$

becomes the Aalen estimate

$$\hat{\beta}_n(t) = \int_0^t \frac{J_n(s)}{Y_n(s)} dN_n(s) = \sum \frac{D_j}{Y_n(X_j)},$$

where D_j is the indicator of death for the j th individual. The corresponding convolution-smoothed estimate is given by

$$\begin{aligned} \tilde{\beta}_n(t) &= \int_0^T W\left(\left(\frac{t-u}{b_n}\right) \frac{J_n(u)}{b_n}\right) dN_n(u) \\ &= \sum_{j=1}^n W\left(\frac{t-X_j}{b_n}\right) \frac{D_j}{Y_n(X_j)}. \end{aligned}$$

Now $Y_n(s)$ is binomially distributed with parameters n and $[1 - F(s)][1 - G(s-)]$ (Aalen, 1976, Lemma 4.2). By the Glivenko-Cantelli theorem, the assumption (i) in Theorem 2.5 is satisfied for t such that $\sigma(t) > 0$, where $\sigma(s) = [1 - F(s)][1 - G(s-)]$.

Remark. The problem of the choice of window size may be attacked various ways. As in the case of ordinary density estimation, one may derive an asymptotically optimal window or choose a window which minimizes some error criterion such as the average squared error or the integrated square error. See Ramlau-Hansen (1983).

Acknowledgements

It is a pleasure to acknowledge the many helpful discussions with my supervisor Professor M.E. Thompson during the preparation of this paper. I greatly appreciated her encouragement and constructive criticism. Thanks are due also to the editor and the referee for their helpful suggestions.

References

- [1] O.O. Aalen, Nonparametric inference for a family of counting processes. *Ann. Statist.* 6 (1978) 701-726.
- [2] S.J. Bean and C.P. Tsokos, Developments in nonparametric density estimation, *Int. Statist. Rev.* 48 (1980) 267-287.

- [3] R.L. Dykstra and P.A. Laud, Bayesian nonparametric approach to reliability, *Ann. Statist.* 9 (1981) 356-367.
- [4] R.D. Gill, *Censoring and Stochastic Integrals* (Mathematisch Centrum Amsterdam, 1980).
- [5] R.Z. Hasminskii and I.A. Ibragimov, *Statistical Estimation, Asymptotic Theory* (Springer-Verlag, New York, 1981).
- [6] Y.P. Mack, Remarks on some smoothed empirical distribution functions and processes, *Bulletin of Informatics and Cybernetics* 21 (1984) 29-35.
- [7] W. Nelson, Theory and application of hazard plotting for censored failure data, *Technometrics* 14 (1972) 945-965.
- [8] H. Ramlau-Hansen, Smoothing counting process intensities by means of kernel functions, *Ann. Statist.* 11 (1983) 453-466.
- [9] M. Rosenblatt, Remarks on some nonparametric estimates of a density function, *Ann. Math. Statist.* 27 (1956) 832-837.