Journal of Mathematics and Modeling in Finance (JMMF)
Vol. 3, No. 1, Winter & Spring 2023
Research paper

Estimation of the hazard rate function in the presence of measurement errors

Parviz Nasiri[1], Roghaieh Kheirazar[2], Abbas Rasouli[3], Ali Shadrokh[4]

[1] University of Payam Noor, Tehran, Iran
pnasiri45@yahoo.com

[2] University of Payam Noor, Tehran, Iran
r.kheirazar65@gmail.com

[3] Department of Statistics, Faculty of Sciences, University of Zanjan, Zanjan, Iran
rasouli@znu.ac.ir

[4] University of Payam Noor, Tehran, Iran
ali_shadrokh@yahoo.com

Abstract:
In this article, according to the importance of the hazard rate function criterion in the evaluation of statistical distributions, its estimation methods are presented. Here, we suggest estimators for the hazard rate function. First, we use the standard deconvolution kernel density estimator and suggest a plug-in estimator. In the following we investigate asymptotic behavior of our estimator. For another estimator, we construct the new estimation the hazard rate function according plug-in and CDF. Finally, we consider the performance of the suggested estimators by simulation.

*Keywords:*Hazard Rate Function, Additive Measurement Errors, Standard Deconvolution Kernel Density estimator, Mean Square Error, Local Polynomial Estimator.

1 Introduction

The hazard rate function is one of the important criteria in statistics. This criterion plays an important role in examining the flexibility of distributions according to the type of location, scale and shape parameters in statistical distributions.

The hazard function is defined as the event rate at time t conditional on survival until time t or later. Let T be a lifetime random variable with continuous distribution function F, survival function $\bar{F} = 1 - F$ and density function f. The hazard rate function is defined as

$$\lambda(t) = \lim_{\Delta \to 0} \frac{P(t \leq T < t + \Delta)}{\Delta.\bar{F}(t)} = \frac{f(t)}{1 - F(t)}, \tag{1}$$

[2]Corresponding author

Received: 02/03/2023 Accepted: 26/06/2023
https://doi.org/10.22054/JMMF.2023.72868.1084

In statistical data analysis, it is important to provide the appropriate distribution for the goodness of fit. For this purpose, it is important to estimate the density function or distribution function in order to use the risk rate function criterion according to equation (1). For more information, refer to articles [3] and [17].

Our second proposed estimator is obtained by using first suggested estimator and the local polynomial regression. A detailed discussion of the local polynomial estimator, can be found in [6]. The local linear principle was first put to use in the hazard setting in the works of [10], [18] and [11]. Nielsen and Tanggaard [12] developed a general class of hazard rate estimates based on local fitting and under a counting processes framework which facilitates broad patterns of data, not necessarily i.i.d. and quite general censoring mechanisms. Mammen and Nielsen [8] extended the method in the multi-dimensional case while estimating old age mortality with a calendar effect. Nielsen et al. [13] developed the density version of the work of [12] and identified the need for development of suitable for practical use bandwidth selectors. See also [2] and the references therein for applications in reliability and medicine related problems. Refer to the [1] for full history of application of local polynomial regression in estimating hazard rate function. All of the sources mentioned, estimate the hazard rate function without considering the measurement error, but we want to fit the local polynomial in hazard rate function with considering measurement error.

In estimating the hazard rate function, since most of the time, instead of actual data or observations, accompanying observations, including risk, $Y_i - \varepsilon_i$ are available. If the actual observations can be considered as

$$Y_i = X_i + \varepsilon_i \quad , \quad i = 1, 2, ..., n \tag{2}$$

Regarding the classical approach to this model, it is assumed that the probability distribution of ε is exactly known; although in many real-life situations this condition cannot be justified. However, in most practical applications, we are able to estimate the error density function of ε from replicated measurements (cf. [9], p. 88.). The well-known deconvolution estimator of f_X is the standard deconvolution kernel density estimator

$$\begin{aligned}\hat{f}_X(x) &= \frac{1}{2\pi}\int \exp(-itx)\phi_K(tb)\,\frac{\frac{1}{n}\sum_{j=1}^{n}\exp(itY_j)}{\phi_\varepsilon(t)}dt \\ &= \frac{1}{n}\sum_{j=1}^{n}\frac{1}{2\pi}\int e^{it(Y_j - x)}\frac{\phi_K(tb)}{\phi_\varepsilon(t)}dt,\end{aligned} \tag{3}$$

where $K : \mathbb{R} \to \mathbb{R}^+$ is kernel function, $b > 0$ a bandwidth parameter and ϕ_K, ϕ_ε are characteristic functions of K and ε, respectively.

2 Plug-in Estimation

In this section, to estimate the hazard rate function parameter according to the model

$$Y_i = X_i + \varepsilon_i \quad , \quad i = 1, 2, ..., n \tag{4}$$

We consider the following multiplicative model.

$$Z_i = T_i \delta_i \quad , \quad i = 1, 2, ..., n \tag{5}$$

such that

$$\ln(Z_i) = \ln(T_i) + \ln(\delta_i) \quad , \quad i = 1, 2, ..., n \tag{6}$$

where $X_i = \ln(T_i),\ Y_i = \ln(Z_i),\ \varepsilon_i = \ln(\delta_i)$. It is easy to show that the estimator of $\hat{f}(t)$ is

$$\hat{f}(t) = |J|\hat{f}(x) = \frac{1}{t}(\ln(t)) \tag{7}$$

Now based on the equation (7) the estimator of $\hat{\lambda}(t)$ is given by

$$\hat{\lambda}(t) = \frac{\hat{f}(x)}{\int_t^{\tau_F} \hat{f}(x)dx} = \frac{\hat{f}(\ln(t))}{t[1 - \hat{F}_X(\ln(t))]}, \tag{8}$$

where $\hat{F}_X(t) = \int_{\nu_F}^t \hat{f}_X(x)dx$ and $\nu_F =: \sup\{x : F(x) = 0\}$.

3 Local Polynomial Estimator

Let $x_1, ..., x_m$ are m selected points from support of data. For select these point on a sub set $[0, \tau_{F_X}]$ with $\tau_{F_X} := \inf\{x : F_X(x) = 1\}$, we partition the interval into m disjoint subintervals $\{I_i, i = 1, ..., m\}$ of equal length $\Delta = \tau_{F_Y}/m$, and let the x_i be the center of each bin I_i. Explicitly, the bin centers can be defined as $x_i = (i - 0.5)\Delta,\ i = 1, ..., m$. The most natural way to construct an empirical estimate of the hazard rate is to divide the relative frequency by the empirical survival function. So empirical estimator for the hazard rate function can be defined as

$$\hat{\lambda}(x_i) = \frac{f_i}{\Delta\left(n - \sum_{j=1}^{i-1} f_j + 1\right)}, \qquad i = 1, 2, ..., m, \tag{9}$$

where f_i is the number of observations in the interval I_i, i.e. is the frequency of the ith bin. This estimator has been employed by [14] and [15] to estimate the failure rate function.

Suppose we want to estimate the hazard rate function at a given point x_0. The local polynomial estimator of order p approximates the function λ by a pth order polynomial $\hat{\lambda}(z) \simeq \sum_{j=0}^{p} \beta_j (z-x_0)^j$, where the local parameters $\beta = (\beta_0, ..., \beta_k)$ are fitted locally by a weighted least squares regression problem, via minimization of

$$\sum_{i=1}^{m} \left[\hat{\lambda}(x_i) - \sum_{j=0}^{p} \beta_j (x_i - x_0)^j \right]^2 K_b(x_i - x_0),$$

where $K_b(.) = b^{-1}K(./b)$, K a kernel function and b is the bandwidth. The local polynomial estimator $\hat{\beta} = (\hat{\beta}_0, ..., \hat{\beta}_p)^T$ with $\hat{\beta}_j = \hat{\lambda}^{(j)}(x_0; p)/j!$ is

$$\hat{\beta} = \left(\mathbf{X}^T \mathbf{W} \mathbf{X}\right)^{-1} \mathbf{X}^T \mathbf{W} \Lambda, \tag{10}$$

so the estimator for the $\lambda(x_0)$ of order p, is given by $\hat{\beta}_0$, i.e.

$$\hat{\lambda}(x_0) = \mathbf{e}_1^T \left(\mathbf{X}^T \mathbf{W} \mathbf{X}\right)^{-1} \mathbf{X}^T \mathbf{W} \Lambda, \tag{11}$$

where

$$\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}_{(p+1)\times 1}, \quad \mathbf{X} = \begin{pmatrix} 1 & (x_1 - x_0) & \cdots & (x_1 - x_0)^p \\ 1 & (x_2 - x_0) & \cdots & (x_2 - x_0)^p \\ \vdots & \vdots & \ddots & \vdots \\ 1 & (x_m - x_0) & \cdots & (x_m - x_0)^p \end{pmatrix}_{m\times(p+1)},$$

$$\Lambda = \begin{pmatrix} \hat{\lambda}(x_1) \\ \hat{\lambda}(x_2) \\ \vdots \\ \hat{\lambda}(x_m) \end{pmatrix}_{m\times 1},$$

$\mathbf{W} = diag(K_b(x_i - x_0))_{m\times m}$ and $\hat{\lambda}(x_i)$, $i = 1, ..., m$ are (8) estimators related to m selected points.

Now when we work with contaminated data, with using our suggested estimator (8) and local polynomial regression, we suggest our second estimator for hazard rate function.

For contaminated data, Let $z_1, ..., z_m$ are centers of m bins $\{I_i, i = 1, ..., m\}$ instead of the above center points $x_1, ..., x_m$, according to additive model (2), i.e. $z_i = x_i + \varepsilon_i$. Suppose $\hat{\lambda}(z_i)$, $i = 1, ..., m$ are (8) estimators related to these points, i.e. at any selected point z_i, we with using $Y_1, ..., Y_n$ estimate hazard rate by our suggested plug-in estimator (8) introduced in the previous section. Our goal is suggestion new estimator for hazard rate at the point x_0, like the estimator (11). But we need to make adjustments and replacing the unobserved quantities $(x_i - x_0)$ and $K_b(x_i - x_0)$.

For $p = 0$, a rate-optimal estimator has been developed by [7]. Their technique is similar to the one employed in density deconvolution problems studied in [16], see also [3]. It consists of replacing the unobserved $K_b(x_i - x_0)$ by an observable quantity $L^*(z_i - x_0)$ satisfying

$$E\left(L^*\left(Z_i - x_0\right)|X_i\right) = K_b\left(X_i - x_0\right),$$

where $L^*(u) = L_1(-u)$ and $L_1(z) = \frac{1}{2\pi}\int e^{-itz}\frac{\phi_k(tb)}{\phi_v(t)}dt$. Following this idea and from [5], our new estimator will be as follows

$$\hat{\beta} = \left(\mathbf{Z}^T\mathbf{L}\mathbf{Z}\right)^{-1}\mathbf{Z}^T\mathbf{L}\Lambda, \tag{12}$$

and our new suggested estimator for the $\lambda(x)$ of order p, is given by $\hat{\beta}_0$, i.e.

$$\hat{\lambda}(x_0; p) = \mathbf{e}_1^T\left(\mathbf{Z}^T\mathbf{L}\mathbf{Z}\right)^{-1}\mathbf{Z}^T\mathbf{L}\Lambda, \tag{13}$$

where

$$\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}_{(p+1)\times 1}, \ \mathbf{Z} = \begin{pmatrix} 1 & (z_1 - x_0) & \cdots & (z_1 - x_0)^p \\ 1 & (z_2 - x_0) & \cdots & (z_2 - x_0)^p \\ \vdots & \vdots & \ddots & \vdots \\ 1 & (z_m - x_0) & \cdots & (z_m - x_0)^p \end{pmatrix}_{m\times(p+1)},$$

$$\Lambda = \begin{pmatrix} \hat{\lambda}\left(z_1\right) \\ \hat{\lambda}\left(z_2\right) \\ \vdots \\ \hat{\lambda}\left(z_m\right) \end{pmatrix}_{m\times 1},$$

$\mathbf{L} = diag(L_b^*\left(z_i - x_0\right))_{m\times m}$ and $\hat{\lambda}(z_i)$, $i = 1, ..., m$ are (8) estimators related to m selected points.

4 Asymptotic Results for the Local Polynomial Estimator

In this section we derive the asymptotical bias and variance of the local polynomial estimator $\hat{\lambda}(t, p)$. The bias and variance of the estimator (12) are derived immediately from its definition:

$$\begin{aligned} E\left(\hat{\beta}\right) &= \left(\mathbf{Z}^T\mathbf{L}\mathbf{Z}\right)^{-1}\mathbf{Z}^T\mathbf{L}\mathbf{m} \\ &= \beta + \left(\mathbf{Z}^T\mathbf{L}\mathbf{Z}\right)^{-1}\mathbf{Z}^T\mathbf{L}\mathbf{r} \\ Var\left(\hat{\beta}\right) &= \left(\mathbf{Z}^T\mathbf{L}\mathbf{Z}\right)^{-1}\left(\mathbf{Z}^T\Sigma\mathbf{Z}\right)\left(\mathbf{Z}^T\mathbf{L}\mathbf{Z}\right)^{-1}, \end{aligned} \tag{14}$$

where $\mathbf{m} = \{E\hat{\lambda}(z_1), ..., E\hat{\lambda}(z_m)\}^T$, $\mathbf{r} = \mathbf{m} - \mathbf{Z}\beta$ and $\Sigma = diag(L_b^{*2}(z_i - x_0)var(\hat{\lambda}(z_i)))$

These exact bias and variance expressions are not diretctly usable, since they depend on unknown quantities. So there is a need for approximating bias and variance. The following notation will be used

$$\begin{aligned} &\mu_j = \textstyle\int u^j L^*(u)\,du, \\ &\upsilon_j = \textstyle\int u^j L^{*2}(u)\,du \\ &\mathbf{S} = (\mu_{j+\ell})_{(p+1)\times(p+1)}, \quad j,\ell = 0,1,...,p \\ &\tilde{\mathbf{S}} = (\mu_{j+\ell+1})_{(p+1)\times(p+1)}, \quad j,\ell = 0,1,...,p \\ &\mathbf{S}^* = (\upsilon_{j+\ell})_{(p+1)\times(p+1)}, \quad j,\ell = 0,1,...,p \\ &\mathbf{C}_p = (\mu_{p+1}, ..., \mu_{2p+1})^T \\ &\tilde{\mathbf{C}}_p = (\mu_{p+2}, ..., \mu_{2p+2})^T. \end{aligned}$$

We assume that the kernel K is a symmetric probability density function with compact support.

Theorem 4.1. *Assume that $f_Y(x_0) > 0$, $\tau_{F_Y} := \inf\{y : F_Y(y) = 1\} < \infty$, $b \to 0$ and $mb \to \infty$ as $m \to \infty$. Then the asymptotic varaince of (13) is given by*

$$var\left\{\hat{\lambda}(x_0; p)\right\} = \boldsymbol{e}_1^T \boldsymbol{S}^{-1}\boldsymbol{S}^*\boldsymbol{S}^{-1}\boldsymbol{e}_1 \frac{\sigma^2(x_0)}{f_Y(x_0)\,nmb} + o_P\left(\frac{1}{nmb}\right), \tag{15}$$

The proof of this theorem is referred to the Appendix. For selection p, the order of local polynomial estimator, notice a given bandwidth b, a large value of p would expectedly reduce the modeling bias, but would cause a large variance and a considerable computational cost. A detailed discussion of this topic, can be found in [6]. Finally, they recommend the use of the lowest odd order, i.e. $p = 1$, or occasionally $p = 3$.

Regarding the error distribution, the kernel function, minimizing the approximated MISE, gives the rule-of-thumb optimum bandwidth value as the following:

$$b_{ROT,N} = \left(\frac{4}{\gamma}\right)^{1/\beta}(\log n)^{-1/\beta} = \sqrt{2}\sigma_\varepsilon(\log n)^{-1/2}.$$

In the case of the homoscedastic Laplacian errors (ordinary smooth), the rule-of-thumb bandwidth becomes

$$b_{ROT,L} = \left(\frac{5\sigma_\varepsilon^4}{n}\right)^{1/9}$$

where σ_ε^2 is the variance of the measurement error. In the R package `decon` the function `bw`.dnrd applies the rule of thumb methods to get the bandwidth value.

5 Estimation of $\lambda(t)$ using CDF estimator

In this section we present estimation the hazard rate function according plug-in and CDF. For this propose we consider $\hat{\lambda}(t)$ as:

$$\hat{\lambda}_c(x_0) = \frac{\hat{f}(x_0)}{1-\hat{F}_\lambda(x_0)} = \frac{1}{1-\hat{F}_\lambda(x_0)}\left[\frac{1}{n}\sum_{j=1}^{n}\frac{1}{2\pi}\int e^{it(y_j-x_0)}\frac{\phi_k(tb)}{\phi_\varepsilon(t)}dt\right] \tag{16}$$

where

$$\hat{F}_\lambda\left(x_0\right) = \frac{1}{2} - \frac{1}{n}\sum_{j=1}^{n}\frac{1}{\pi}\int_0^\lambda \frac{1}{\omega}\mathrm{Im}\left\{\frac{e^{i\omega(Y_j-x_0)}}{\phi_\varepsilon\left(\omega\right)}\right\}d\omega \tag{17}$$

which introduced by [4] and Im{.} stands for the imaginary part.

6 Simulation study and conclusion

To illustrate our estimators, we present here estimation of $\lambda(t)$ based on estimators $\hat{\lambda}(t,p)$, $\hat{\lambda}(t)$ and $\hat{\lambda}_c(t)$. For the measurement error we consider the normal distribution with zero mean and variance σ_ε^2 and three measurement error distribution scenarios:
1- Weibull distribution with shape parameter 2 and scale parameter 1.
2- Gamma distribution with shape parameter 3 and scale parameter $\sqrt{3}$.
3- Lindley distribution with shape parameter 3 and scale parameter 1.2.

The simulation has been repeated for 1000 observation samples of size $n = 50, 200, 500, 1000$. For the local polynomial estimator, we take $m = 25$ and according to recommendation, $p = 1$. Tables 1-6 and figures 1-3 summarize the empirical mean square error and values of local polynomial estimators $\hat{\lambda}(t,p)$, $\hat{\lambda}(t)$ and plug-in estimators $\hat{\lambda}_c(t)$ results of different set ups. The results in tables 1-6 show that with increasing samples size, the mean square error of estimators decreasing. In the case of normal errors with zero expectation with changing 15% to 30% contamination the mean square errors of estimators in tables 1-6 decreeing.

Table 1: Mean square error of estimators $\hat{\lambda}(t,p)$, $\hat{\lambda}(t)$ and $\hat{\lambda}_c(t)$ of **Weibull** distribution based on Local polynomial and plug-in estimators when Normal error with zero exception and 15% contamination.

		t				
		0.25	0.5	0.75	1.00	1.25
n=50	$\hat{\lambda}(t,p)$	0.080	0.101	0.191	0.362	0.576
	$\hat{\lambda}(t)$	0.193	0.206	0.300	0.461	0.684
	$\hat{\lambda}_C(t)$	0.193	0.207	0.302	0.474	0.790
n=200	$\hat{\lambda}(t,p)$	0.031	0.035	0.063	0.109	0.208
	$\hat{\lambda}(t)$	0.057	0.074	0.097	0.140	0.246
	$\hat{\lambda}_C(t)$	0.057	0.074	0.097	0.143	0.254
n=500	$\hat{\lambda}(t,p)$	0.012	0.014	0.022	0.039	0.080
	$\hat{\lambda}(t)$	0.022	0.027	0.035	0.052	0.095
	$\hat{\lambda}_C(t)$	0.023	0.028	0.038	0.059	0.101
n=1000	$\hat{\lambda}(t,p)$	0.009	0.009	0.016	0.026	0.048
	$\hat{\lambda}(t)$	0.014	0.015	0.0232	0.033	0.055
	$\hat{\lambda}_C(t)$	0.014	0.016	0.025	0.038	0.065

Table 2: Mean square error of estimators $\hat{\lambda}(t,p)$, $\hat{\lambda}(t)$ and $\hat{\lambda}_c(t)$ of **Weibull** distribution based on Local polynomial and plug-in estimators when Normal error with zero exception and 30% contamination.

		t				
		0.25	0.5	0.75	1.00	1.25
n=50	$\hat{\lambda}(t,p)$	0.024	0.048	0.091	0.186	0.323
	$\hat{\lambda}(t)$	0.090	0.102	0.142	0.235	0.360
	$\hat{\lambda}_C(t)$	0.090	0.102	0.143	0.260	0.443
n=200	$\hat{\lambda}(t,p)$	0.007	0.015	0.030	0.053	0.118
	$\hat{\lambda}(t)$	0.030	0.033	0.044	0.067	0.136
	$\hat{\lambda}_C(t)$	0.030	0.033	0.046	0.074	0.180
n=500	$\hat{\lambda}(t,p)$	0.003	0.006	0.012	0.023	0.038
	$\hat{\lambda}(t)$	0.011	0.012	0.018	0.029	0.047
	$\hat{\lambda}_C(t)$	0.012	0.014	0.021	0.035	0.081
n=1000	$\hat{\lambda}(t,p)$	0.002	0.003	0.007	0.015	0.037
	$\hat{\lambda}(t)$	0.005	0.006	0.009	0.018	0.040
	$\hat{\lambda}_C(t)$	0.006	0.008	0.014	0.034	0.079

Table 3: Mean square error of estimators $\hat{\lambda}(t,p)$, $\hat{\lambda}(t)$ and $\hat{\lambda}_c(t)$ of **gamma** distribution based on Local polynomial and plug-in estimators when Normal error with zero exception and 15% contamination.

		t				
		0.5	1.0	1.5	2.0	2.5
n=50	$\hat{\lambda}(t,p)$	0.018	0.039	0.059	0.085	0.150
	$\hat{\lambda}(t)$	0.034	0.053	0.068	0.090	0.156
	$\hat{\lambda}_C(t)$	0.034	0.053	0.069	0.092	0.161
n=200	$\hat{\lambda}(t,p)$	0.006	0.009	0.017	0.026	0.039
	$\hat{\lambda}(t)$	0.012	0.013	0.020	0.028	0.041
	$\hat{\lambda}_C(t)$	0.012	0.013	0.020	0.030	0.044
n=500	$\hat{\lambda}(t,p)$	0.001	0.003	0.006	0.008	0.015
	$\hat{\lambda}(t)$	0.004	0.006	0.008	0.010	0.017
	$\hat{\lambda}_C(t)$	0.004	0.007	0.011	0.015	0.025
n=1000	$\hat{\lambda}(t,p)$	0.000	0.001	0.003	0.005	0.009
	$\hat{\lambda}(t)$	0.002	0.003	0.004	0.006	0.010
	$\hat{\lambda}_C(t)$	0.003	0.004	0.006	0.008	0.013

Table 4: Mean square error of estimators $\hat{\lambda}(t,p)$, $\hat{\lambda}(t)$ and $\hat{\lambda}_c(t)$ of **gamma** distribution based on Local polynomial and plug-in estimators when Normal error with zero exception and 30% contamination.

		t				
		0.5	1.0	1.5	2.0	2.5
n=50	$\hat{\lambda}(t,p)$	0.008	0.018	0.025	0.038	0.055
	$\hat{\lambda}(t)$	0.019	0.025	0.028	0.041	0.059
	$\hat{\lambda}_C(t)$	0.019	0.025	0.030	0.052	0.103
n=200	$\hat{\lambda}(t,p)$	0.002	0.005	0.008	0.012	0.018
	$\hat{\lambda}(t)$	0.005	0.007	0.009	0.013	0.020
	$\hat{\lambda}_C(t)$	0.005	0.007	0.010	0.017	0.038
n=500	$\hat{\lambda}(t,p)$	0.001	0.002	0.003	0.005	0.006
	$\hat{\lambda}(t)$	0.002	0.002	0.003	0.006	0.007
	$\hat{\lambda}_C(t)$	0.002	0.003	0.006	0.014	0.036
n=1000	$\hat{\lambda}(t,p)$	0.001	0.001	0.002	0.004	0.005
	$\hat{\lambda}(t)$	0.002	0.002	0.003	0.005	0.006
	$\hat{\lambda}_C(t)$	0.002	0.003	0.005	0.009	0.014

Table 5: Mean square error of estimators $\hat{\lambda}(t,p)$, $\hat{\lambda}(t)$ and $\hat{\lambda}_c(t)$ of **Lindley** distribution based on Local polynomial and plug-in estimators when Normal error with zero exception and 15% contamination.

		t				
		0.25	0.75	1.00	1.25	1.50
n=50	$\hat{\lambda}(t,p)$	0.011	0.137	0.225	0.388	0.512
	$\hat{\lambda}(t)$	0.024	0.219	0.290	0.437	0.563
	$\hat{\lambda}_C(t)$	0.024	0.219	0.292	0.450	0.580
n=200	$\hat{\lambda}(t,p)$	0.004	0.043	0.080	0.102	0.147
	$\hat{\lambda}(t)$	0.009	0.066	0.105	0.122	0.165
	$\hat{\lambda}_C(t)$	0.009	0.066	0.106	0.124	0.170
n=500	$\hat{\lambda}(t,p)$	0.001	0.012	0.022	0.037	0.051
	$\hat{\lambda}(t)$	0.003	0.021	0.029	0.042	0.056
	$\hat{\lambda}_C(t)$	0.003	0.023	0.033	0.049	0.066
n=1000	$\hat{\lambda}(t,p)$	0.001	0.009	0.017	0.025	0.038
	$\hat{\lambda}(t)$	0.001	0.012	0.021	0.028	0.040
	$\hat{\lambda}_C(t)$	0.002	0.014	0.022	0.035	0.046

Table 6: Mean square error of estimators $\hat{\lambda}(t,p)$, $\hat{\lambda}(t)$ and $\hat{\lambda}_c(t)$ of **Lindley** distribution based on Local polynomial and plug-in estimators when Normal error with zero exception and 30% contamination

		t				
		0.25	0.75	1.00	1.25	1.50
n=50	$\hat{\lambda}(t,p)$	0.012	0.068	0.109	0.159	0.266
	$\hat{\lambda}(t)$	0.015	0.101	0.144	0.190	0.297
	$\hat{\lambda}_C(t)$	0.016	0.105	0.149	0.200	0.360
n=200	$\hat{\lambda}(t,p)$	0.003	0.020	0.035	0.049	0.072
	$\hat{\lambda}(t)$	0.004	0.028	0.042	0.058	0.080
	$\hat{\lambda}_C(t)$	0.004	0.030	0.047	0.066	0.103
n=500	$\hat{\lambda}(t,p)$	0.001	0.009	0.014	0.017	0.021
	$\hat{\lambda}(t)$	0.001	0.012	0.017	0.022	0.029
	$\hat{\lambda}_C(t)$	0.001	0.013	0.021	0.029	0.045
n=1000	$\hat{\lambda}(t,p)$	0.001	0.005	0.009	0.013	0.018
	$\hat{\lambda}(t)$	0.001	0.007	0.010	0.015	0.020
	$\hat{\lambda}_C(t)$	0.001	0.009	0.014	0.026	0.040

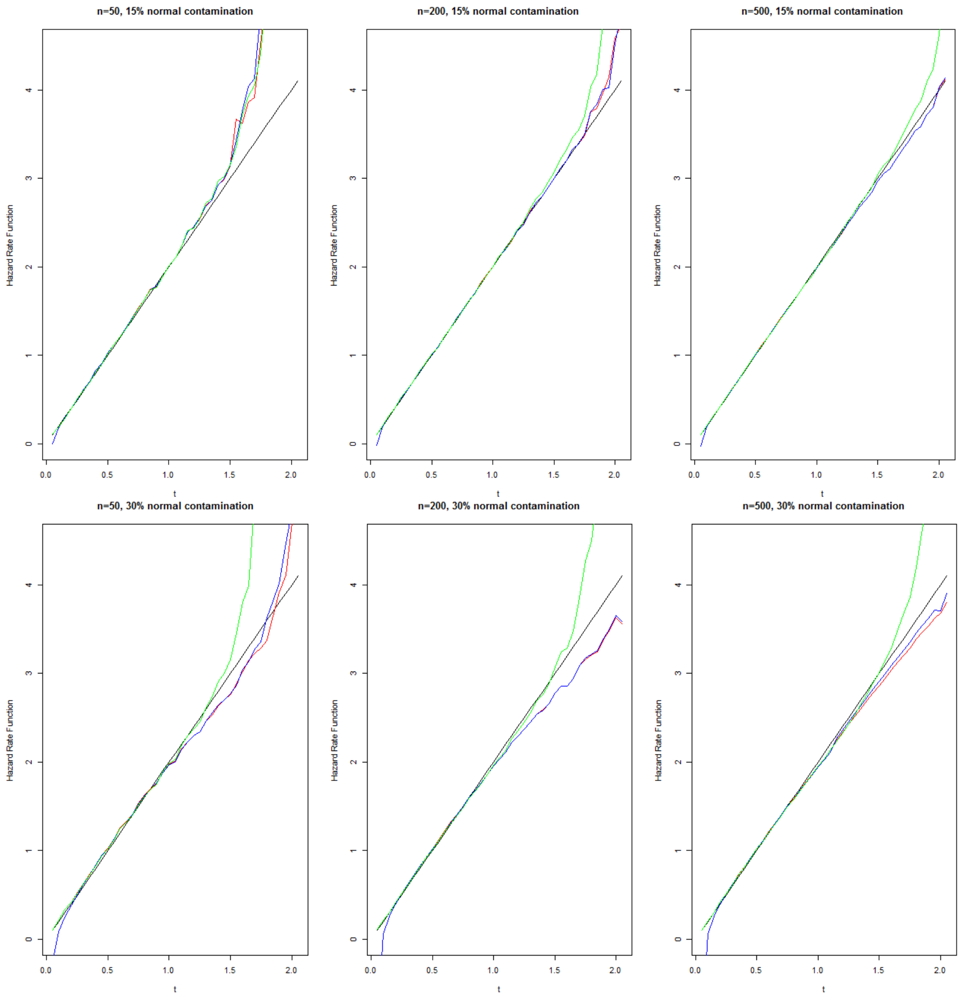

Figure 1: Estimating $\lambda(t)$ when actual data have **weibull** distribution with shape 2 and scale 1 with normal contamination. black line: Actual value, Blue line: Estimator (8), Red line: Estimator (13), green line: Estimator (16)

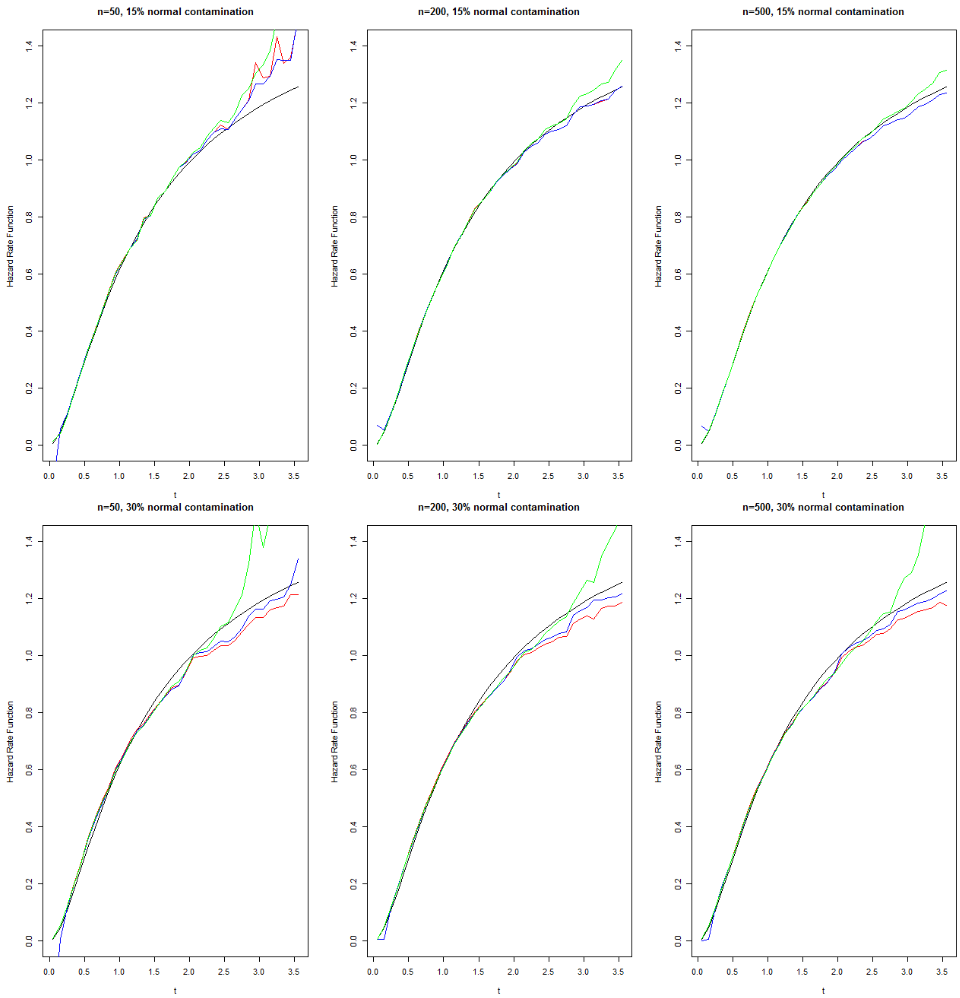

Figure 2: Estimating $\lambda(t)$ when actual data have **gamma** distribution with shape 3 and scale $\sqrt{3}$ with normal contamination. black line: Actual value, Blue line: Estimator (8), Red line: Estimator (13), green line: Estimator (16)

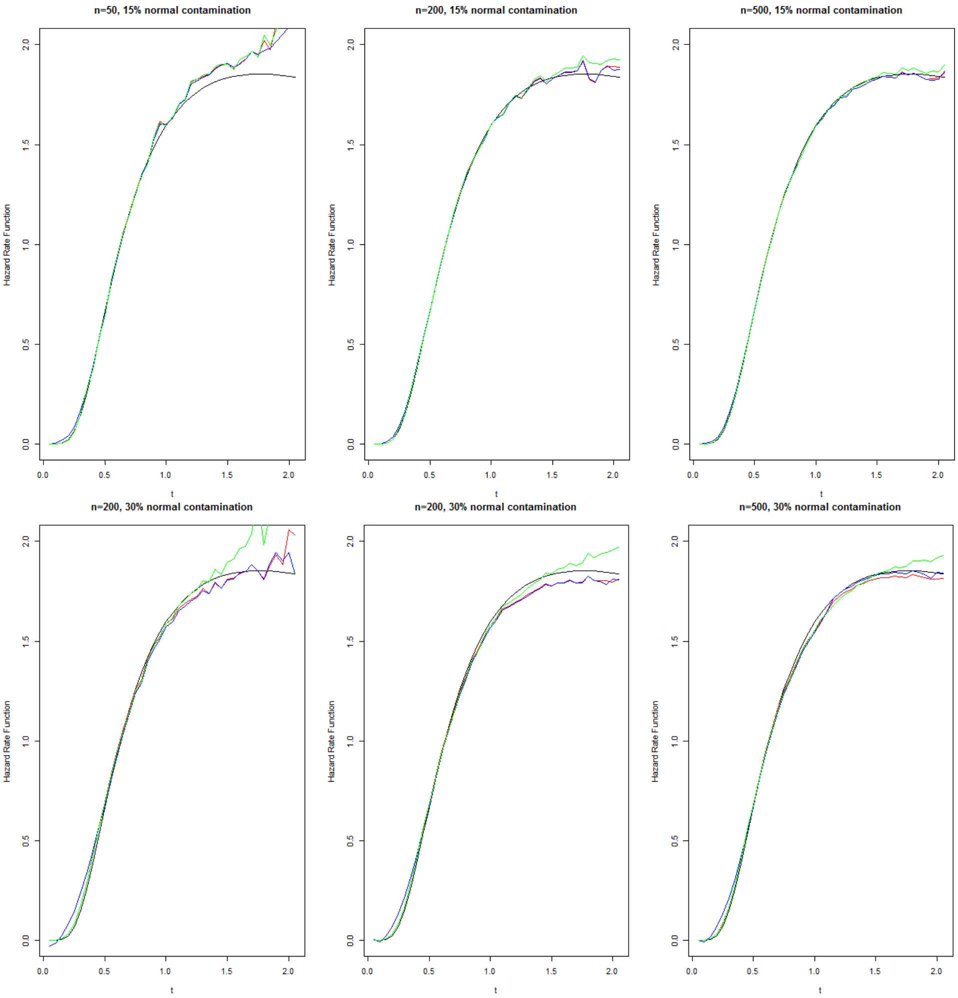

Figure 3: Estimating $\lambda(t)$ when actual data have **Lindley** distribution with shape 3 and scale 1.2 with normal contamination. black line: Actual value, Blue line: Estimator (8), Red line: Estimator (13), green line: Estimator (16)

Acknowledgement

We thank the chief editor, the associate editor and two anonymous reviewers for insightful comments that helped clarify and improve the paper significantly.

Bibliography

[1] D. Bagkavos, *Local linear hazard rate estimation and bandwidth selection*, Annals of the Institute of Statistical Mathematics, 63 (2011), pp. 1019-1046.

[2] D. Bagkavos, P. N. Patil, *Local polynomial fitting in failure rate estimation*, IEEE Transactions on Reliability, 56 (2008), pp. 126-163.

[3] R. J. Carroll, P. Hall, *Optimal rates of convergence for deconvolving a density*, Journal of American Statistics Association, 83 (1988), pp. 1184-1186.

[4] I. Dattner, A. Goldenshluger, A. Juditsky, *On deconvolution of distribution functions*, The Annals of Statisitcs, 39 (2011), pp. 2477-2501.

[5] A. Delaigle, J. Fan, R. J. Carroll, *A Design-Adaptive Local Polynomial Estimator for the Errors-in-Variables Problem*, Journal of the American Statistical Association, 104 (2009), pp. 348-359.

[6] J. Fan, I. Gijbels, *Local Polynomial Modeling and Its Applications*, Chapman and Hall, London, 1996.

[7] J. Fan, Y. K. Truong, *Nonparametric regression with errors in variables*, The Annals of Statistics, 21 (1993), pp. 1900-1925.

[8] E. Mammen, J. P. Nielsen, *A general approach to the predictability issue in survival analyses*, Biometrika, 94 (1998), pp. 873-892.

[9] A. Meister, *Deconvolution Problems in Nonparametric Statistics*, Springer, Berlin, Germany, 2009.

[10] H. G. Müller, J. L. Wang, W. B. Capra, *From lifetables to hazard rates: The transformation approach*, Biometrika, 84 (1997), pp. 881-892.

[11] J. P. Nielsen, *Marker dependent hazard estimation from local linear estimation*, Scandinavian Actuarial Journal, 2 (1998), pp. 113-124.

[12] J. P. Nielsen, C. G. Tanggaard, *Boundary and bias correction in kernel hazard estimation*, Scandinavian Journal of Statistics, 28 (2001), pp. 675-698.

[13] J. P. Nielsen, C. Tanggaard, C. Jones, *Local linear density estimation for filtered survival data*, Statistics, 42 (2009), pp. 167-186.

[14] G. Pulcini, *Modeling the failure data of a repairable equipment with bathtub type failure intensity*, Reliability Engineering and System Safety, 71 (2001), pp. 209-218.

[15] B, Rai, N. Singh, *Customer-rush near warranty expiration limit and nonparametric hazard rate estimation from known mileage accumulation rates*, IEEE Transactions on Reliability, 55 (2006), pp. 480-489.

[16] L. A, Stefanski, R. J. Carroll, *Deconvoluting Kernel Density Estimators*, Statistics: A Journal of Theoretical and Applied Statistics, 21 (1990), pp. 169-184.

[17] M. P. Wand, *Data-based choice of histogram bin width*, The American Statistician, 51 (1997), pp. 59-64.

[18] J. L. Wang, H. G. Müller, W. B. Capra, *Analysis of oldest-old mortality: Lifetables revisited*, Annals of Statitics, 28 (1998), pp. 126-163.

Appendix

Proof of Theorem 4.1 . Here, for convenience we use $f(.)$ instead of $f_Y(.)$. With notation

$$\mathbf{S}_m = \mathbf{Z}^T \mathbf{L} \mathbf{Z} = \left(S_{m,j+\ell}\right)_{(p+1)\times(p+1)}$$
$$S_{m,j} = \sum_{i=1}^{m} (z_i - x_0)^j L_b^* (z_i - x_0), \qquad j, \ell = 0, 1, ..., p$$
$$\mathbf{S}_m^* = \mathbf{Z}^T \Sigma \mathbf{Z} = \left(S^*_{m,j+\ell}\right)_{(p+1)\times(p+1)}$$
$$S^*_{m,j} = \sum_{i=1}^{m} (z_i - x_0)^j L_b^{*2} (z_i - x_0) \operatorname{var}\left(\hat{\lambda}(z_i)\right), \qquad j, \ell = 0, 1, ..., p,$$

the variance (14) can be re-expressed as

$$Var\left(\hat{\beta}\right) = \mathbf{S}_m^{-1} \mathbf{S}_m^* \mathbf{S}_m^{-1}.$$

Now we find approximations for the $\mathbf{S}_m$ and $\mathbf{S}_m^*$. For elements of $\mathbf{S}_m$ we have

$$S_{m,j} = ES_{m,j} + O_P\left\{\sqrt{Var\left(S_{m,j}\right)}\right\},$$

$$\begin{aligned} ES_{m,j} &= mE\left\{(Z_i - x_0)^j L_b^* (Z_i - x_0)\right\} \\ &= m\int (y - x_0)^j \frac{1}{b} L^*\left(\frac{y - x_0}{b}\right) f(y)\, dy \\ &= mb^j \int u^j L^*(u) f(x_0 + hu)\, du \\ &= mb^j \int u^j L^*(u) \{f(x_0) + o(1)\}\, du, \qquad \textit{provided } b \to 0 \\ &= mb^j \{f(x_0)\mu_j + o(1)\} \end{aligned}$$

and

$$
\begin{aligned}
O_P\left\{\sqrt{Var\left(S_{m,j}\right)}\right\} &= O_P\left\{\sqrt{mVar\left((Z_i-x_0)^j L_b^*\left(Z_i-x_0\right)\right)}\right\} \\
&= O_P\left\{\sqrt{mE\left((Z_i-x_0)^{2j} L_b^{*2}\left(Z_i-x_0\right)\right)}\right\} \\
&= O_P\left\{\sqrt{m\int_0^{\tau_{F_Y}} (y-x_0)^{2j}\frac{1}{b^2}L^{*2}\left(\frac{y-x_0}{b}\right)f\left(y\right)dy}\right\} \\
&= O_P\left\{\sqrt{mb^{2j-1}\int_0^{\tau_{F_Y}} u^{2j}L^{*2}\left(u\right)f\left(x_0+bu\right)du}\right\} \\
&= O_P\left\{\sqrt{mb^{2j-1}\left\{\underbrace{f\left(x_0\right)\int_0^{\tau_{F_Y}} u^{2j}L^{*2}\left(u\right)du+o\left(1\right)}_{O(1)}\right\}}\right\}, \quad \textit{provided } b\to 0 \\
&= O_P\left\{\sqrt{mb^{2j-1}}\right\} \\
&= mb^j\left\{O_P\left(1/\sqrt{mb}\right)\right\}.
\end{aligned}
$$

$$
\begin{aligned}
S_{m,j} &= ES_{m,j} + O_P\left\{\sqrt{Var\left(S_{m,j}\right)}\right\} \\
&= mb^j\left\{f\left(x_0\right)\mu_j + o\left(1\right) + O_P\left(1/\sqrt{mb}\right)\right\} \\
&= mb^j f\left(x_0\right)\mu_j\left\{1+o_P\left(1\right)\right\},
\end{aligned}
$$

provided $mb\to\infty$ as $m\to\infty$. From this we have

$$
\mathbf{S}_m = mf\left(x_0\right)\mathbf{BSB}\left\{1+o_P\left(1\right)\right\},
$$

where $\mathbf{B}=diag(1,b,...,b^p)$. Using similar arguments for elements of $\mathbf{S}_m^*$ we have

$$
S_{m,j}^* = ES_{m,j}^* + O_P\left\{\sqrt{Var\left(S_{m,j}^*\right)}\right\}
$$

and

$$
\begin{aligned}
ES_{m,j}^* &= mE\left\{(Z_i-x_0)^j L_b^{*2}\left(Z_i-x_0\right)\mathrm{var}\left(\hat{\lambda}\left(Z_i\right)\right)\right\} \\
&= m\int (y-x_0)^j\frac{1}{b^2}L^{*2}\left(\frac{y-x_0}{b}\right)\mathrm{var}\left(\hat{\lambda}\left(y\right)\right)f\left(y\right)dy \\
&= mb^{j-1}\int u^j L^{*2}\left(u\right)\mathrm{var}\left(\hat{\lambda}\left(x_0+bu\right)\right)f\left(x_0+bu\right)du \\
&= \frac{mb^{j-1}}{n}\left\{f\left(x_0\right)\sigma^2\left(x_0\right)\int u^j L^{*2}\left(u\right)du + o\left(1\right)\right\}, \qquad \textit{provided } \ b\to 0 \\
&= \frac{mb^{j-1}}{n}\left\{f\left(x_0\right)\sigma^2\left(x_0\right)\upsilon_j + o\left(1\right)\right\}.
\end{aligned}
$$

$$
\begin{aligned}
O_P\left\{\sqrt{Var\left(S_{m,j}^{*}\right)}\right\} &= O_P\left\{\sqrt{mVar\left((Z_i-x_0)^j L_b^{*2}\left(Z_i-x_0\right)\text{var}\left(\hat{\lambda}\left(Z_i\right)\right)\right)}\right\}\\
&= O_P\left\{\sqrt{mE\left((Z_i-x_0)^{2j} L_b^{*4}\left(Z_i-x_0\right)\text{var}^2\left(\hat{\lambda}\left(Z_i\right)\right)\right)}\right\}\\
&= O_P\left\{\sqrt{m\int_0^{\tau_{F_Y}}\left(y-x_0\right)^{2j}\frac{1}{b^4}L^{*4}\left(\frac{y-x_0}{b}\right)\text{var}^2\left(\hat{\lambda}\left(y\right)\right)f\left(y\right)dy}\right\}\\
&= O_P\left(\sqrt{mb^{2j-3}\int_0^{\tau_{F_Y}}u^{2j}L^{*4}\left(u\right)\text{var}^2\left(\hat{\lambda}\left(x_0+bu\right)\right)f\left(x_0+bu\right)du}\right)\\
&= O_P\left(\sqrt{\frac{mb^{2j-3}}{n^2}\left\{\underbrace{\sigma^4\left(x_0\right)f\left(x_0\right)\int_0^{\tau_{F_Y}}u^{2j}L^{*4}\left(u\right)du+o\left(1\right)}_{O(1)}\right\}}\right),\\
&\quad \textit{provided}\ \ b\to 0\\
&= O_P\left(\sqrt{\frac{mb^{2j-3}}{n^2}}\right)\\
&= \frac{mb^{j-1}}{n}\left\{O_P\left(1/\sqrt{mb}\right)\right\}.
\end{aligned}
$$

By inserting

$$
\begin{aligned}
S_{m,j}^{*} &= ES_{m,j}^{*}+O_P\left\{\sqrt{Var\left(S_{m,j}^{*}\right)}\right\}\\
&= \frac{mb^{j-1}}{n}\left\{f\left(x_0\right)\sigma^2\left(x_0\right)\upsilon_j+o\left(1\right)+O_P\left(1/\sqrt{mb}\right)\right\}\\
&= \frac{mb^{j-1}}{n}f\left(x_0\right)\sigma^2\left(x_0\right)\upsilon_j\left\{1+o_P\left(1\right)\right\},
\end{aligned}
$$

provided $mb\to\infty$ as $m\to\infty$. So

$$
\mathbf{S}_m^{*}=\frac{m}{nb}f\left(x_0\right)\sigma^2\left(x_0\right)\mathbf{BS}^{*}\mathbf{B}\left\{1+o_P\left(1\right)\right\}.
$$

we have

$$
\begin{aligned}
\text{var}\left(\hat{\beta}\right) &= \mathbf{S}_m^{-1}\mathbf{S}^{*}\mathbf{S}_m^{-1}\\
&= \left(mf\left(x_0\right)\mathbf{BSB}\left\{1+o_P\left(1\right)\right\}\right)^{-1}\left(\frac{m}{nb}f\left(x_0\right)\sigma^2\left(x_0\right)\mathbf{BS}^{*}\mathbf{B}\left\{1+o_P\left(1\right)\right\}\right)\\
&\qquad\times\left(mf\left(x_0\right)\mathbf{BSB}\left\{1+o_P\left(1\right)\right\}\right)^{-1}\\
&= \frac{\sigma^2\left(x_0\right)}{f\left(x_0\right)nmb}\mathbf{B}^{-1}\mathbf{S}^{-1}\mathbf{S}^{*}\mathbf{S}^{-1}\mathbf{B}^{-1}\left\{1+o_P\left(1\right)\right\}
\end{aligned}
$$

and since $\hat{\lambda}(x_0;p)=\hat{\beta}_0=\mathbf{e}_1^T\hat{\beta}$, so

$$\begin{aligned}
\operatorname{var}\left\{\hat{\lambda}\left(x_0;p\right)\right\} &= \operatorname{var}\left(\mathbf{e}_1^T\hat{\beta}\right)\\
&= \frac{\sigma^2\left(x_0\right)}{f\left(x_0\right)nmb}\mathbf{e}_1^T\mathbf{B}^{-1}\mathbf{S}^{-1}\mathbf{S}^*\mathbf{S}^{-1}\mathbf{B}^{-1}\mathbf{e}_1\left\{1+o_P\left(1\right)\right\}\\
&= \frac{\sigma^2\left(x_0\right)}{f\left(x_0\right)nmb}\mathbf{e}_1^T\mathbf{S}^{-1}\mathbf{S}^*\mathbf{S}^{-1}\mathbf{e}_1\left\{1+o_P\left(1\right)\right\}\\
&= \mathbf{e}_1^T\mathbf{S}^{-1}\mathbf{S}^*\mathbf{S}^{-1}\mathbf{e}_1\frac{\sigma^2\left(x_0\right)}{f\left(x_0\right)nmb}+o_P\left(\frac{1}{nmb}\right).
\end{aligned}$$

How to Cite: Parviz Nasiri[1], Roghaieh Kheirazar[2], Abbas Rasouli[3], Ali Shadrokh[4], *Estimation of the hazard rate function in the presence of measurement errors*, Journal of Mathematics and Modeling in Finance (JMMF), Vol. 3, No. 1, Pages:49–66, (2023).