Electronic Journal of Applied Statistical Analysis
EJASA, Electron. J. App. Stat. Anal.
http://siba-ese.unisalento.it/index.php/ejasa/index
e-ISSN: 2070-5948
DOI: 10.1285/i20705948v10n1p1

A comparison study between three different Kernel estimators for the Hazard rate function

By Salha, Rasheed

Published: 26 April 2017

Electronic Journal of Applied Statistical Analysis
Vol. 10, Issue 01, April 2017, 1-13
DOI: 10.1285/i20705948v10n1p1

A comparison study between three different Kernel estimators for the Hazard rate function

Raid B. Salha*and Ali J. Rasheed

The Islamic University, Palestinian Territory, Occupied
Department of Mathematics

Published: 26 April 2017

Kernel estimation is one of the most important data analytical tool, if we consider the non parametric approach in the estimation of the probability density function. In parallel, it is used to estimate the hazard rate function, which is one of the most important ways for representing the life time distribution in the survival analysis. As the support of the hazard rate function is in the non negative part of the real line $[0, \infty)$, its will be under the boundary effect near zero when the estimation is done using symmetric kernels such as the Gaussian kernel. Two kernel estimators for the hazard rate function were proposed using asymmetric kernels are the Reciprocal Inverse Gaussian and Inverse Gaussian kernel estimators to avoid the high bias near zero. In this paper, we conduct a theoretical comparison between those estimators by looking at their asymptotic bias, variance and the mean squared error. Also, a comparison of the practical performance of the three estimators based on simulated and real data will be present.

keywords: Inverse Gaussian kernel, reciprocal inverse Gaussian kernel, asymptotic bias, asymptotic variance, mean squared error.

1 Introduction

The hazard rate function is the instantaneous failure rate or (a force of morality), its represent the probability of failure between time x and $x + \Delta$, given that there were

*Corresponding author: rbsalha@iugaza.edu.ps

ISSN: 2070-5948
http://siba-ese.unisalento.it/index.php/ejasa/index

no failures up to time x. For more details see Cox and Oakes (1984). The hazard rate function is defined as follows:

Definition 1.1. *The hazard rate function or age-specific failure rate, defined by:*

$$r(x) = \lim_{\Delta \longrightarrow 0} \frac{P(x < X \leq x + \Delta | x \leq X)}{\Delta} \tag{1.1}$$

and by the definition of the conditional probability, we have

$$r(x) = \frac{f(x)}{1 - F(x)} \tag{1.2}$$

where $f(x)$ is the pdf of the distribution and $F(x)$ is the cdf.

Let $X_1, X_2, \cdots, X_n$ a random sample from a distribution with an unknown probability density function f satisfying the following assumptions:

(A_1) The unknown density function $f(x)$ has a continuous second derivative $f^{(2)}(x)$.

(A_2) The bandwidth $h = h_n$ is a sequence of positive numbers and satisfies $h \to 0$ and $nh \to \infty$ as $n \to \infty$.

(A_3) The kernel K is a bounded probability density function of order 2 and symmetric about the zero.

Watson and Leadbetter (1964) has proposed the following kernel estimator for the hazard rate function.

Definition 1.2. *The kernel estimator for the hazard rate function with bandwidth h is given by:*

$$\hat{r}(x) = \frac{\hat{f}(x)}{1 - \hat{F}(x)} \tag{1.3}$$

where,
$\hat{f}(x) = \frac{1}{nh} \sum_{i=1}^{n} K\left(\frac{x - X_i}{h}\right)$, $\hat{F}(x) = \frac{1}{nh} \sum_{i=1}^{n} \int_0^x K\left(\frac{x - X_i}{h}\right)$ *and K is bounded symmetric kernel function with $\int_0^\infty K(u)du = 1$.*

Definition 1.3. *The Gaussian kernel estimator for the hazard rate function $\hat{r}_G(x)$ is defined by:*

$$\hat{r}_G(x) = \frac{\hat{f}_G(x)}{\hat{S}_G(x)} \tag{1.4}$$

where, $\hat{f}_G(x)$ and $\hat{S}_G(x)$ are defined as follows:
$\hat{f}_G(x) = \frac{1}{nh} \sum_{i=1}^{n} K_G\left(\frac{x - X_i}{h}\right)$, $\hat{S}_G(x) = 1 - \hat{F}_G(x) = 1 - \frac{1}{nh} \sum_{i=1}^{n} \int_0^x K_G\left(\frac{u - X_i}{h}\right) du$,
where $K_G(x)$ is given by $K_G(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \forall x \in \Re$.

Note that the support of the hazard rate function is in the non-negative part of the real line $[0, \infty)$, so when the estimation is based on symmetric kernels it will be under the **boundary effect** (called **a boundary bias** problem) near the zero. its causes that the estimator of the hazard rate function will take values outside the support.
To solve this problem, Chen (2000) has replaced the symmetric kernels by asymmetric Gamma kernel, which never assigns weight outside the support. Scaillet (2004) used this idea and proposed two new classes of density estimators, rely on the use of inverse Gaussian (IG) and the reciprocal inverse Gaussian (RIG) kernels in the place of the Gamma kernel. In Salha (2012), the estimation of the hazard rate function using the IG kernel has been considered. In Salha (2013), the estimation of the hazard rate function using the RIG kernel has been considered.
Now, we state the following conditions under which, Scaillet (2004) has proposed the IG and RIG kernel estimators of the pdf.

Conditions

(C_1) Let $X_1, X_2, \cdots, X_n$ be a random sample from a distribution with an unknown probability density function f defined on $[0, \infty)$, such that f is twice continuously differentiable, and $\int_0^\infty (x^3 f''(x))^2 dx < \infty$.

(C_2) h is a smoothing parameter satisfying $h + \frac{1}{nh} \to 0$, and $nh^{\frac{5}{2}} \to 0$, as $n \to \infty$.

Scaillet (2004) has proposed the following IG kernel function as follows:

Definition 1.4. *The IG kernel function is defined by :*

$$K_{IG(x,\frac{1}{h})}(u) = \frac{1}{\sqrt{2\pi h u^3}} exp\left(-\frac{1}{2hx}\left(\frac{u}{x} - 2 + \frac{x}{u}\right)\right), u > 0 \tag{1.5}$$

where $h + \frac{1}{nh} \to 0$ *as* $n \to \infty$.

Using this kernel, Scaillet (2004) proposed the following IG kernel estimators of the pdf and cdf as follows:

Definition 1.5. *The IG kernel estimator of the pdf is defined by :*

$$\hat{f}_{IG}(x) = \frac{1}{n} \sum_{i=1}^{n} K_{IG(x,\frac{1}{h})}(X_i) \tag{1.6}$$

Definition 1.6. *The IG kernel estimator of the cdf is defined by :*

$$\hat{F}_{IG}(x) = \int_0^x \hat{f}_{IG}(u) du = \frac{1}{n} \sum_{i=1}^{n} \int_0^x K_{IG}(u, \frac{1}{h})(X_i) du. \tag{1.7}$$

Using definitions 1.5 and 1.6, we propose the IG kernel estimator for the hazard rate function as follows:

Definition 1.7. *The IG kernel estimator for the hazard rate function is given by :*

$$\hat{r}_{IG}(x) = \frac{\hat{f}_{IG}(x)}{1 - \hat{F}_{IG}(x)}. \tag{1.8}$$

Under the same conditions, Scaillet (2004) has proposed the following RIG kernel function.

Definition 1.8. *The RIG kernel function is defined by :*

$$K_{RIG(\frac{1}{x-h},\frac{1}{h})}(u) = \frac{1}{\sqrt{2\pi hu}} exp\left(-\frac{x-h}{2h}\left(\frac{u}{x-h} - 2 + \frac{x-h}{u}\right)\right), u > 0 \tag{1.9}$$

where $h + \frac{1}{nh} \to 0$ *as* $n \to \infty$ *.*

Using this kernel, Scaillet (2004) proposed the following RIG kernel estimators of the pdf and cdf as follows:

Definition 1.9. *The RIG kernel estimator of the pdf is defined by :*

$$\hat{f}_{RIG}(x) = \frac{1}{n}\sum_{i=1}^{n} K_{RIG(\frac{1}{x-h},\frac{1}{h})}(X_i) \tag{1.10}$$

Definition 1.10. *The RIG kernel estimator of the cdf is defined by :*

$$\hat{F}_{RIG}(x) = \int_0^x \hat{f}_{RIG}(u)du = \frac{1}{n}\sum_{i=1}^{n}\int_0^x K_{RIG}(\frac{1}{u-h}, \frac{1}{h})(X_i)du. \tag{1.11}$$

Using definitions 1.9 and 1.10, we propose the RIG kernel estimator for the hazard rate function as follows:

Definition 1.11. *The RIG kernel estimator for the hazard rate function is given by:*

$$\hat{r}_{RIG}(x) = \frac{\hat{f}_{RIG}(x)}{1 - \hat{F}_{RIG}(x)}. \tag{1.12}$$

Salha (2012) and Salha (2013) studied the asymptotic properties for the estimators in Equations 1.8 and 1.12 such as the asymptotic normality, the strong consistency and investigated the optimal bandwidth that minimizes the mean squared error (MSE) and the asymptotic mean squared error ($AMSE$). In Section 2, we present a theoretical comparison between the three estimators from equations 1.4, 1.8 and 1.12. Also, a practical comparison between the three estimators to test their performance is given in Section 3. Finally, our conclusions will be given in Section 4.

2 A theoretical comparison

In this section, we make a brief theoretical comparison of the $\hat{r}_G(x)$, $\hat{r}_{RIG}(x)$ and $\hat{r}_{IG}(x)$ estimators by looking at their asymptotic bias, variance and the MSEs under the assumptions A_1, A_2 and A_3 for the Gaussian and under the conditions C_1 and C_2 for both RIG and IG. Table 1 summarizes the results for the bias, variance and the MSEs for both $\hat{r}_{RIG}(x)$ and $\hat{r}_{IG}(x)$ are taken from Salha (2012) and Salha (2013).

The Estimator	The Bias	The Variance
$\hat{r}_G(x)$	$\frac{f''(x)h^2}{2S(x)} + o(h)$	$\frac{f(x)}{2nh\sqrt{\pi}S^2(x)} + o(\frac{1}{nh})$
$\hat{r}_{IG}(x)$	$\frac{x^3 f''(x)h}{2S(x)} + o(h)$	$\frac{1}{2n\sqrt{\pi h}} x^{-\frac{3}{2}} \frac{r(x)}{S(x)} + o(n^{-1}h^{-\frac{1}{2}})$
$\hat{r}_{RIG}(x)$	$\frac{x f''(x)h}{2S(x)} + o(h)$	$\frac{1}{2n\sqrt{\pi h}} x^{-\frac{1}{2}} \frac{r(x)}{S(x)} + o(n^{-1}h^{-\frac{1}{2}})$

Table 1: Summary for the Bias and Variance for the three estimators

2.1 The Bias

By looking to those expressions for the Bias in Table 1, we conclude the following remarks:

1. Since the expressions of the $Bias(\hat{r}_{RIG}(x))$ and $Bias(\hat{r}_G(x))$ increases in xh and h^2 respectively, and hence near the zero ($x \in (0,h)$), we have $xh < h^2$, which imply that $Bias(\hat{r}_{RIG}(x)) < Bias(\hat{r}_G(x))$.

2. $Bias(\hat{r}_{RIG}(x))$ and $Bias(\hat{r}_{IG}(x))$ increases in xh and x^3h respectively, and hence for any $(x > 1)$ we have $xh < x^3h$, which imply that $Bias(\hat{r}_{RIG}(x)) < Bias(\hat{r}_{IG}(x))$.

3. If $0 < x < 1$ we have $x^3 < x$ and hence $x^3h < xh < h^2$ which imply that $Bias(\hat{r}_{IG}(x)) < Bias(\hat{r}_{RIG}(x)) < Bias(\hat{r}_G(x))$.

2.2 The Variance

By looking to the expressions for the Variance in Table 1, we conclude the following remarks:

1. The expressions of the $Var(\hat{r}_{RIG}(x))$ and $Var(\hat{r}_G(x))$ decreases in $\sqrt{xh}$ and h respectively, and hence as $x \in (0,h)$ we have $\sqrt{xh} < h$, which imply that $Var(\hat{r}_{RIG}(x)) > Var(\hat{r}_G(x))$.

2. The expressions of the $Var(\hat{r}_{RIG}(x))$ and $Var(\hat{r}_{IG}(x))$ decreases in $\sqrt{xh}$ and $\sqrt{x^3h}$ respectively, and hence as $0 < x < 1$ we have $\sqrt{xh} > \sqrt{x^3h}$, which imply that $Var(\hat{r}_{RIG}(x)) < Var(\hat{r}_G(x))$.

3. The expressions of the $Var(\hat{r}_{RIG}(x))$ and $Var(\hat{r}_{IG}(x))$ decreases in $\sqrt{xh}$ and $\sqrt{x^3h}$ respectively, and hence as $x > 1$ we have $\sqrt{xh} < \sqrt{x^3h}$, which imply that $Var(\hat{r}_{RIG}(x)) > Var(\hat{r}_{IG}(x))$.

2.3 The MSE

By Table 1, the squared bias for the RIG decreases if $x \in (0,h)$ and hence as h decreases we have squared bias for the RIG less than squared bias for Gaussian which imply that $MSE(\hat{r}_{RIG}(x)) < MSE(\hat{r}_G(x))$. A similar result hols for $\hat{r}_{IG}(x)$, we have $MSE(\hat{r}_{IG}(x)) < MSE(\hat{r}_G(x))$.

2.4 The Optimal AMSE

For any $x \in [0, \infty)$, the optimal AMSE ($AMSE^*$) for the three estimators is the same and it is given by

$$AMSE^*(\hat{r}_{RIG}(x)) = AMSE^*(\hat{r}_G(x)) = AMSE^*(\hat{r}_{IG}(x)) = \frac{5}{4}\left(\frac{f(x)}{2\sqrt{\pi}}\right)^{\frac{4}{5}}\left(\frac{n^{-\frac{4}{5}}f''(x)^{\frac{2}{5}}}{S^2(x)}\right).$$

The $AMSE^*(\hat{r}_{IG}(x))$ was driven in Salha (2012) and following the same techniques, we can derive the other two optimal AMSE. Now, we derive the $AMSE^*(\hat{r}_G(x))$.
From Table 1, we have

$$MSE(\hat{r}_G(x)) = \left(\frac{f''(x)h^2}{2S(x)}\right)^2 + \frac{f(x)}{2nh\sqrt{\pi}S^2(x)} + o(h^2) + o(\frac{1}{nh}). \tag{2.1}$$

Then under the conditions on the bandwidth h in C_2 and as $n \to \infty$, the $AMSE$ is given by

$$AMSE(\hat{r}_G(x)) = \left(\frac{f''(x)}{2S(x)}\right)^2 h^4 + \left(\frac{f(x)}{2n\sqrt{\pi}S^2(x)}\right) h^{-1} \tag{2.2}$$

To find $AMSE^*(\hat{r}_G(x))$, the value of the optimal bandwidth, that minimizes $AMSE(\hat{r}_G(x))$, must be substituted in Equation (2.2).
Now, differentiate Equation (2.2) and equating it to zero, we obtain

$$\left(\frac{f''(x)}{S(x)}\right)^2 h^3 - \left(\frac{f(x)}{2n\sqrt{\pi}S^2(x)}\right) h^{-2} = 0 \tag{2.3}$$

Solve Equation (2.3) for h, we obtain the optimal bandwidth as follows:

$$h = \left(\frac{f(x)}{2n\sqrt{\pi}(f''(x))^2}\right)^{\frac{1}{5}} \tag{2.4}$$

Now, substitute the value of the optimal bandwidth from Equation (2.4) into Equation (2.2), we get

$$AMSE^*(\hat{r}_G(x)) = \left(\frac{f''(x)}{2S(x)}\right)^2 \left(\frac{f(x)}{2n\sqrt{\pi}(f''(x))^2}\right)^{\frac{4}{5}} + \left(\frac{f(x)}{2n\sqrt{\pi}S^2(x)}\right)\left(\frac{f(x)}{2n\sqrt{\pi}(f''(x))^2}\right)^{-\frac{1}{5}} \tag{2.5}$$

Now, by simplifying Equation (2.5), the following holds

$$AMSE^*(\hat{r}_G(x)) = \frac{5}{4}\left(\frac{f(x)}{2\sqrt{\pi}}\right)^{\frac{4}{5}}\left(\frac{n^{-\frac{4}{5}}f''(x)^{\frac{2}{5}}}{S^2(x)}\right).$$

3 Practical Comparison

In this section, we compare between the three estimators of the hazard rate function by studying their performance using simulated and real data. The bandwidths has been chosen according to the following rules.
First, for the Gaussian estimator, the following rule from Silverman (1986) has been used to find the bandwidth

$$h^* = 1.06\hat{\sigma} n^{-\frac{1}{5}}, \tag{3.1}$$

where,

$$\hat{\sigma}^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2.$$

and for the RIG and IG estimators, we used the bandwidth selection procedure that has been proposed by Scaillet (2004). This procedure indicates that if lnX follows a normal distribution with parameters μ and σ^2, then the optimal bandwidths for the RIG and IG estimators are given respectively by

$$h^{**} = \left(\frac{16\sigma^5 exp(\frac{1}{8}(-17\sigma^2 + 20\mu))}{12 + 4\sigma^2 + \sigma^4}\right)^{\frac{2}{5}} n^{-\frac{2}{5}}, \tag{3.2}$$

and

$$h^{***} = \left(\frac{16\sigma^5 exp(\frac{1}{8}(17\sigma^2 - 20\mu))}{12 + 68\sigma^2 + 225\sigma^4}\right)^{\frac{2}{5}} n^{-\frac{2}{5}}, \tag{3.3}$$

where, the unknown parameters σ and μ are estimated as follows:

1. $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} lnx_i$,
2. $\hat{\sigma}^2 = \frac{1}{n-1}\sum_{i=1}^{n}(lnx_i - \bar{x})^2$.

We use the S-Plus program in the implementation of those applications.

3.1 Simulation Studies

In this subsection, the performance of the three estimators are tested using simulation data from the exponential and normal distributions.

3.1.1 Simulation Study 1

A sample of size 200 from the exponential distribution with pdf $f(x) = e^{-x}$ is simulated. After that the density function and the hazard rate functions were estimated using the RIG, IG and the Gaussian estimators. The estimated values and the true functions are plotted in Figure 1 and Figure 2, respectively. The two figures show that the performance of the RIG and IG estimator is better than that of the Gaussian estimator at the boundary near the zero. In the interior the behavior of the three estimators becomes more similar as we get away from the zero. Also the MSE for the hazard rate estimators are listed in Table 2. Table 2 indicates that the RIG estimator has the smallest MSE.

MSE for the Estimators		
$MSE(\hat{r}_{RIG}(x))$	$MSE(\hat{r}_{IG}(x))$	$MSE(\hat{r}_{G}(x))$
0.4588251	0.4686859	0.4812027

Table 2: MSE of the hazard rate estimators for Simulation Study 1

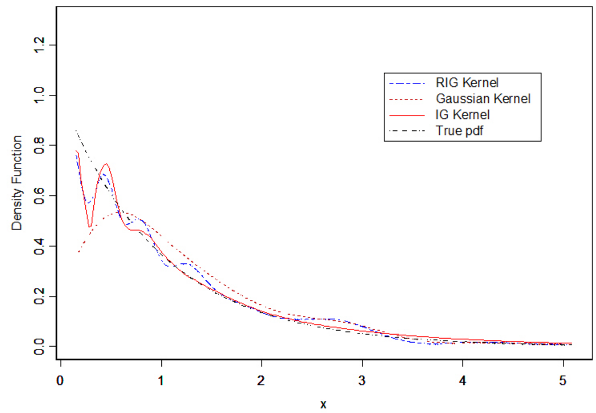

Figure 1: The RIG, IG and Gaussian kernel estimators of the density function for the simulated data of the exponential distribution

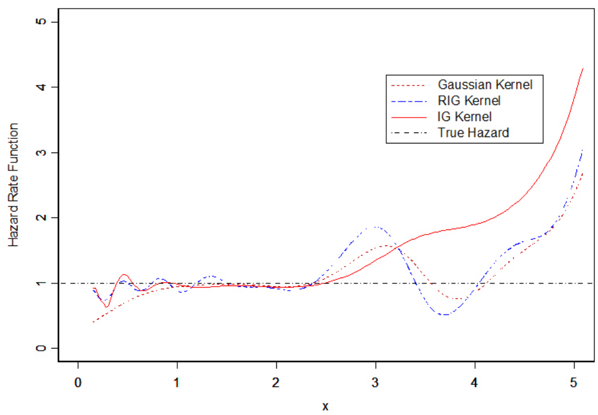

Figure 2: The RIG, IG and Gaussian kernel estimators of the hazard rate function for the simulated data of the exponential distribution

3.1.2 Simulation Study 2

Two samples of size 40 and 100 from the normal distribution with pdf $f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$ are simulated. After that the density function and the hazard rate functions were estimated using the RIG, IG and the Gaussian estimators. The estimated values and the true functions for the sample of size 100 are plotted in Figure 3 and Figure 4, respectively. The two figures show that the performance of the RIG estimator is better than that of the Gaussian and IG estimators at the boundary near the zero. In the interior the behavior of the three estimators becomes more similar as we get away from the zero. Also the MSE for the hazard rate estimators for the two samples are listed in Table 3. Table 3 indicates that the RIG estimator has the smallest MSE. From the results in Table 3, we note the performance of the three estimators for large sample is better than that for the small sample.

MSE **for the Estimators**			
Sample size	$MSE(\hat{r}_{RIG}(x))$	$MSE(\hat{r}_{IG}(x))$	$MSE(\hat{r}_{G}(x))$
40	0.3432449	0.62093	0.6671531
100	0.0412683	0.2949743	0.1256913

Table 3: MSE for the hazard rate estimators for Simulation Study 2

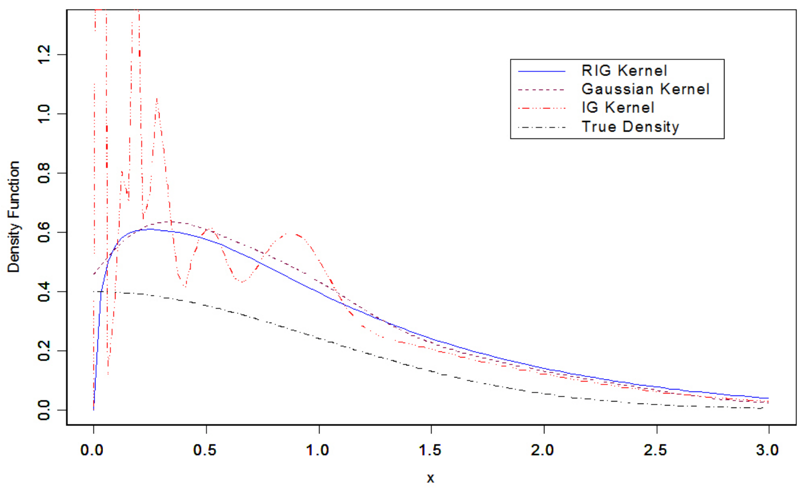

Figure 3: The RIG, IG and Gaussian kernel estimators of the density function for the simulated data of the normal distribution

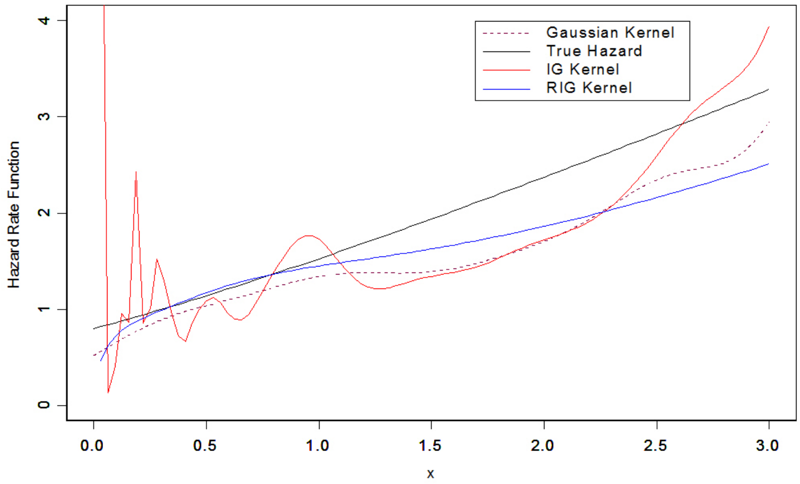

Figure 4: The *RIG*, *IG* and Gaussian kernel estimators of the hazard rate function for the simulated data of the normal distribution

3.2 Real Data

We use the survival time of the lung cancer patients given in data from a study of lung cancer patients conducted by the North Central Cancer Treatment Group, see Loprinzi et al. (1994), to exhibit and compare the practical performances of the Gaussian, *RIG* and *RIG* estimators. We exclude the **censored data** (means some individuals may not observed for the full time to failure which for example stay alive at the end of the study or may leave the study before they die), so here we assume that the applications done using a complete study (without censoring). The data gives the lengths of the treatment spells (in days) of control patients were hospitalized. The objective is to estimate the hazard rate function which in this case represents the instant potential per unit of time that an individual die within the time interval $(x, x+\Delta)$ given that it was known to be alive up to time x.

Figure 5 and Figure 6 show the estimators of the probability density and hazard rate function, respectively. Although the suggested values of the density and hazard rate functions from the estimators are different, they suggest a similar structure for the estimated functions. As we see, the divergence of the estimators gets large at the boundary near the zero and becomes smaller in the interior especially from approximately $x \geq 200$.

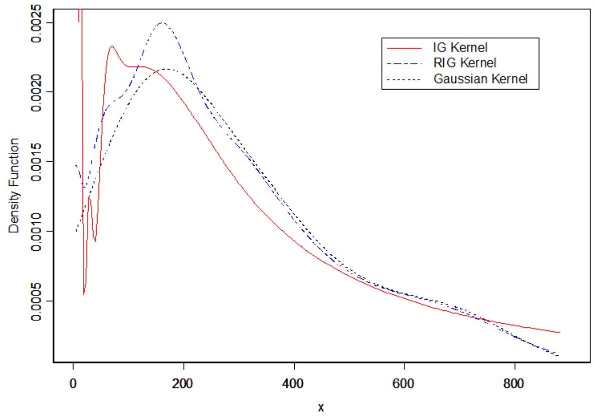

Figure 5: The RIG, IG and Gaussian kernel estimators of the density function for the survival time of the lung cancer patients

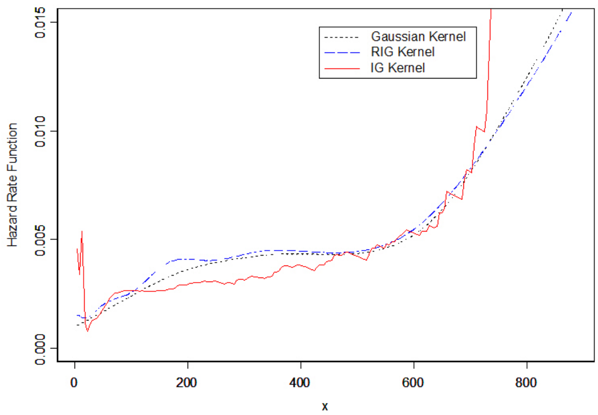

Figure 6: The RIG,IG and Gaussian kernel estimators of the hazard rate function for the survival time of the lung cancer patients

4 Conclusion

In this paper, we compared between three kernel estimators for the hazard rate function, the RIG, IG and the Gaussian estimators. A theoretical comparison between those estimators based on comparing the asymptotic biases, variances and mean squared error indicated that the two asymmetric kernel estimators ($\hat{r}_{IG}$ and $\hat{r}_{RIG}$) are better than the Gaussian at the boundary near the zero. This result leads to deduce that the mean squared errors for both ($\hat{r}_{IG}$ and $\hat{r}_{RIG}$) are less than that of the Gaussian kernel estimator ($\hat{r}_G$) because its based on symmetric kernel. Also, the practical comparison between the three estimators using simulation studies and real data indicated that the performance of the asymmetric kernel estimators ($\hat{r}_{IG}$ and $\hat{r}_{RIG}$)is better than that of $\hat{r}_G$, especially near the zero, confirming the previous theoretical discussion.

In this paper, we presented three estimators of the hazard rate function and compared between them theoretically and practically. The main idea was to replace the symmetric kernel function by asymmetric kernel functions to avoid the boundary bias problem near the zero, when estimating the hazard rate function. To increase the accuracy of these estimators, we suggest to use a variable bandwidth that depends on the points where we estimate the hazard rate function. This variable bandwidth together with the asymmetric kernel will increase the performs of the kernel estimators of the hazard rate function.

References

Chen, S. X. (2000). Probability density function estimation using gamma kernels. *Annals of the Institute of Statistical Mathematics*, 52(3):471–480.

Cox and Oakes (1984). *Analysis of survival data.* Champion and Hall.

Loprinzi, C. L., Laurie, J. A., Wieand, H. S., Krook, J. E., Novotny, P. J., Kugler, J. W., Bartel, J., Law, M., Bateman, M., and Klatt, N. E. (1994). Prospective evaluation of prognostic variables from patient-completed questionnaires. north central cancer treatment group. *Journal of Clinical Oncology*, 12(3):601–607.

Salha, R. (2012). Hazard rate function estimation using inverse gaussian kernel. *The Islamic university of Gaza journal of Natural and Engineering studiesl*, 20(1):73–84.

Salha, R. (2013). Estimating the density and hazard rate functions using reciprocal inverse gaussian kernel. In *Proceedings of the 15th international conference of Applied Stochastic Models and Data Analysis (ASMADA)*, pages 759–766. ASMADA.

Scaillet, O. (2004). Density estimation using inverse and reciprocal inverse gaussian kernels. *Communications in Statistics Theory and Methods*, 16:217–226.

Silverman, B. W. (1986). *Density estimation for statistics and data analysis*, volume 26. CRC press.

Watson, G. and Leadbetter, M. (1964). Hazard analysis i. *Biometrika*, 51:175–184.