Austrian Journal of Statistics
2026, Volume 55, 139–156.
http://www.ajs.or.at/
doi:10.17713/ajs.v55i1.2128

Gamma and Reciprocal Inverse Gaussian Kernel Estimators for Stress Strength Reliability Model

Meriem Medjider
University of Bejaia, Faculty of Exact Sciences
Applied Mathematics Laboratory
06000 Bejaia, Algeria

Karima Lagha
University of Bejaia
Faculty of Exact Sciences
Research Unit LaMOS
06000 Bejaia, Algeria

Abstract

Several parametric estimation methods have been applied to estimate the reliability of a stress strength model. In this paper, we focus on the nonparametric kernel method and propose two nonparametric estimators based, respectively, on the Gamma and Reciprocal Inverse Gaussian kernels, which are nonnegative kernels, free from boundary bias and achieve the optimal rate of convergence for the mean integrated squared error. We introduce reliability estimators and establish their statistical and asymptotic properties, namely: bias, variance, and mean square error. We also study the selection of the smoothing parameter using the rule-of-thumb and unbiased cross-validation approaches, since it plays an important role in kernel estimation. Finally, a simulation study and an application to real data are carried out to highlight the performance of the proposed estimators.

Keywords: stress strength model, nonparametric estimation, Gamma kernel, Reciprocal Inverse Gaussian kernel, bandwidth parameter.

1. Introduction

The stress strength (SST) model is used primarily in reliability engineering, but also in psychology, economics and medicine. In the context of reliability, the SST model compares strength and stress on a system or component, both of which are considered as separate random variables. In their reference work on SST models, Kotz, Lumelskii, and Pensky (2003) detail numerous examples of stress strength models in a survey of the scientific literature. These include applications such as Rocket Engines reliability, Earthquake Resistance and a medical study.

The stress experienced by a component is often represented by the random variable Y and the strength of the component available to overcome the stress is represented by the random variable X. The system fails if $Y > X$; it is a situation in which the stress applied to it is greater than its strength. We can define reliability as the probability that a component will not fail. Model reliability is therefore defined as the probability $R = P(X > Y)$, which can be considered as a measure of component performance. Typically, R is referred to as SST reliability, as it shows the relationship between stress and component strength. It plays an important role in many practical fields, including medicine, quality control and engineering.

Many studies have estimated SST models using various parametric distributions due to their flexibility and ability to model diverse data behaviors. For instance, Kundu and Gupta (2006) used the Weibull distribution, while Raqab, Madi, and Kundu (2008) extended the model using exponentiated distributions. Rezaei, Tahmasbi, and Mahmoodi (2010) investigated the generalized pareto distribution. Asgharzadeh, Valiollahi, and Raqab (2013) introduced the generalized logistic distribution. More recent developments, Jia, Nadarajah, and Guo (2017) addressed the Weibull distributions with arbitrary parameters, and Muhammad, Wang, Li, Yan, and Chang (2020) investigated models under the poisson half logistic distribution. Almarashi, Algarni, and Nassar (2020) and Abu El Azm, Almetwally, Alghamdi, Aljohani, Muse, and Abo-Kasem (2021) extended the methodology using Weibull and exponentiated inverted Weibull distributions, respectively. The maximum likelihood and Bayesian approaches are the most common procedures used to estimate the reliability of these models. In all these works, the variables X and Y have the same probability density function with different parameters. However, in many practical situations, parametric life models may perform poorly due to incorrect assumptions or limited flexibility. A good alternative to the parametric approach is the nonparametric kernel method.

Nonparametric kernel is a method that was first introduced to estimate the probability density function (PDF) (see Parzen (1962)). Over time, it has been extended and improved, particularly for non-negative data, through the development of asymmetric kernel functions (see Chen (1999, 2000), Scaillet (2004), Jin and Kawczak (2003) and Marchant, Bertin, Leiva, and Saulo (2013)). This method has found several applications in different fields, particularly in reliability to estimate the reliability and the failure rate functions (see Bouezmarni, El Ghouch, and Mesfioui (2011), Salha (2013) and Chekkal, Lagha, and Zougab (2023)). Despite its flexibility and minimal assumptions, kernel estimation has not yet been exploited in the reliability analysis of SST models. It offers a valuable advantage in that it does not require prior knowledge on the form of the stress and strength densitys to be estimated, and is easy to implement in practice.

The aim of this paper is to estimate the reliability of the SST model, without any restrictions on density, using the asymmetric kernel method with non-negative support instead of a symmetric kernel to eliminate the problem of high bias in the boundary region, the so-called boundary effect. Notably, Bouezmarni and Scaillet (2005) proposed a Gamma kernel density estimator specifically designed for non-negative data. Their approach effectively reduces boundary bias and demonstrates the advantages of asymmetric kernels in modeling data with support on $[0, \infty)$. Similarly, Hagmann and Scaillet (2007) introduced local multiplicative bias correction techniques that enhance the accuracy of asymmetric estimators. Gustafsson, Hagmann, Nielsen, and Scaillet (2009) further contributed to the field by proposing local transformation kernel estimators for loss distributions, leveraging asymmetric kernels to improve fit in skewed contexts. Their contribution is particularly valuable when estimating the densities of stress and strength. Incorporating such nonparametric techniques opens new possibilities for modeling complex reliability systems without imposing strict distributional assumptions. The current work builds on this idea by evaluating the performance of Gamma (GA) and Reciprocal Inverse Gaussian (RIG) kernels in the estimation of the SST model. These kernel density estimators achieve the best mean integrated squared error (MISE) convergence rate among the class of non-negative kernel density estimators; that they are free of boundary bias and their variance reduces the position at which smoothing is performed away from the boundary. We therefore propose in this paper two nonparametric reliability kernel estimators for the SST model. This work represents a first study in this field and may lead to further studies concerning the choice of kernel and smoothing parameter in order to improve the results obtained by the proposed estimators.

This paper is organized as follows: In Section 2, we briefly review the GA and RIG kernel estimators of PDF. In Section 3, we present the nonparametric reliability estimator of the

SST model based on GA and RIG kernels (denoted by SST-GA and SST-RIG, respectively). We develop their statistical and asymptotic properties (bias, variance and mean squared error (MSE)). In addition, we select the optimal bandwidth with rule-of-thumb (RT) and unbiased cross-validation (UCV) techniques. Sections 4 and 5 present respectively a simulation study and application on real data and we conclude the paper in Section 6. All the proofs are presented in Appendix (Section A).

2. Review on PDF estimation using GA and RIG kernels

Let X be a positive and continuous random variable (r.v.) with unknown PDF f defined on $[0, \infty)$. Let $X_1, ..., X_n$ be a random sample on X, continuous, independent and identically distributed (i.i.d).
The nonparametric GA kernel estimator of the PDF is given by Chen (2000):

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

where $h > 0$ is a bandwidth (or smoothing parameter) and $K_{GA(x,h)}(.)$ is GA kernel defined by

$$K_{GA(x,h)}(t) = \frac{t^{x/h} e^{-t/h}}{h^{x/h+1} \Gamma(x/h+1)}, \quad t > 0, \tag{2}$$

with $\Gamma(.)$ is Gamma function defined as $\Gamma(a) = \int_0^{+\infty} x^{a-1} e^{-x} \, dx, a > 0$.

The nonparametric RIG kernel estimator of the PDF is given by Scaillet (2004):

$$\hat{f}_{RIG}(x) = \frac{1}{n} \sum_{i=1}^{n} K_{RIG(x,h)}\left(X_i\right), \tag{3}$$

where $h > 0$ is a bandwidth (or smoothing parameter) and $K_{RIG(x,h)}(.)$ is the RIG kernel defined by

$$K_{RIG(x,h)}(t) = \frac{1}{\sqrt{2\pi h t}} exp\left[-\frac{x-h}{2h}\left(\frac{t}{x-h} - 2 + \frac{x-h}{t}\right)\right], \; t > 0. \tag{4}$$

Consider the two following conditions:

C1 The function f is twice differentiable and its second derivative is continuous and bounded,

C2 The bandwidth $h = h_n$ satisfies $\lim_{n\to\infty} h = 0$ and $\lim_{n\to\infty} n h^{1/2} = +\infty$.

The expressions of the bias and variance for $\hat{f}_{GA}$ and $\hat{f}_{RIG}$ are respectively derived by Chen (2000) and Scaillet (2004), under the conditions **C1** and **C2** as:

$$\text{Bias}(\hat{f}_{GA}(x)) = h\{f'(x) + \frac{1}{2} x f''(x)\} + o(h), \tag{5}$$

$$\text{Bias}(\hat{f}_{RIG}(x)) = \frac{1}{2} x f''(x) h + o(h) \tag{6}$$

and

$$\text{Var}(\hat{f}_{GA}(x)) = \text{Var}(\hat{f}_{RIG}(x)) = \frac{1}{2\sqrt{\pi}} h^{-1/2} n^{-1} x^{-1/2} f(x) + o(n^{-1} h^{-1/2}). \tag{7}$$

3. SST reliability kernel estimators

In this section, we present the kernel estimators of reliability SST model based on GA and RIG kernels and we calculate bias, variance and mean square error of the estimators in Theorem 3.1. We also investigate the bandwidth selection for the estimators using RT and UCV techniques.

3.1. SST-GA and SST-RIG estimators

The reliability R of the SST model is defined as the probability $R = P(X > Y)$ that a component fails if the stress Y which is applied to it exceeds the strength X and there is no failure otherwise. In another way, it is the probability that a component will survive as long as the strength exceeds the stress applied to it. It is therefore a performance measurement of the component. X and Y are assumed independent random variables with PDFs f and g, respectively. So, R can be written as

$$R = P(X > Y) = \int_0^{+\infty} \int_0^x f(x)g(y)dydx = \int_0^{+\infty} f(x)G(x)dx,$$

where, $G(x) = \int_0^x g(y)dy$ is the CDF of Y calculated at $x,\ x > 0$.

Let $X_1, X_2, \ldots, X_n$ be a random sample on X from unknown PDF f, and let $Y_1, Y_2, \ldots, Y_m$ be a random sample on Y from unknown PDF g and unknown cumulative distribution function (CDF) G, where n and m are the respective sizes of the two samples. The unknown PDFs f and g can be estimated by the kernel estimators given in (1) and (3), respectively for GA and RIG kernels. In addition, the unknown CDF G can be estimated by

$$\hat{G}_k(x) = \int_0^x \hat{g}_k(y)dy = \int_0^x \frac{1}{m} \sum_{j=1}^{m} K_{k(y,h_m)}(Y_j)dy,$$

where k denote GA or RIG, such as K_{GA} and K_{RIG} are GA and RIG kernels defined respectively in (2) and (4). So, we can propose two kernel estimators for the parameter R, defined for $k = GA$ or $k = RIG$ as

$$\hat{R}_k = \frac{1}{nm} \sum_{i=1}^{n} \sum_{j=1}^{m} \int_0^{+\infty} \int_0^x K_{k(x,h_n)}(X_i) K_{k(y,h_m)}(Y_j)dydx.$$

Therefore, $\hat{R}_{GA}$ and $\hat{R}_{RIG}$ are the SST-GA and SST-RIG estimators expressed respectively by

$$\hat{R}_{GA} = \frac{1}{nm} \sum_{i=1}^{n} \sum_{j=1}^{m} \int_0^{+\infty} \int_0^x \left(\frac{X_i^{x/h_n} e^{-X_i/h_n}}{h_n^{x/h_n+1} \Gamma(x/h_n+1)} \right) \left(\frac{Y_j^{y/h_m} e^{-Y_j/h_m}}{h_m^{y/h_m+1} \Gamma(y/h_m+1)} \right) dydx \quad (8)$$

and,

$$\hat{R}_{RIG} = \frac{1}{nm2\pi} \sum_{i=1}^{n} \sum_{j=1}^{m} \int_0^{+\infty} \int_0^x \frac{1}{\sqrt{h_n h_m X_i Y_j}} exp\left[-\frac{x-h_n}{2h_n} \left(\frac{X_i}{x-h_n} - 2 + \frac{x-h_n}{X_i} \right) - \frac{y-h_m}{2h_m} \left(\frac{Y_j}{y-h_m} - 2 + \frac{y-h_m}{Y_j} \right) \right] dydx. \quad (9)$$

3.2. Statistical properties

The Theorem 3.1 below provides the biases, variances and mean square errors (MSEs) of the estimators $\hat{R}_{GA}$ and $\hat{R}_{RIG}$, defined in formulas (8) and (9) under conditions **C1**, **C2** and the following assumptions

A1. $\int_0^{+\infty} G(x)\{f'(x) + \frac{1}{2}xf''(x)\}dx < \infty$,

A2. $\int_0^{+\infty} f(x)\{g(x) + xg'(x)\}dx < \infty$,

A3. $\int_0^{+\infty} \{f'(x) + \frac{1}{2}xf''(x)\}\{g(x) + xg'(x)\}dx < \infty$,

A4. $\int_0^{+\infty} xG(x)f''(x)dx < \infty$,

A5. $\int_0^{+\infty} f(x)\{xg'(x) - g(x)\}dx < \infty$,

A6. $\int_0^{+\infty} xf''(x)\{xg'(x) - g(x)\}dx < \infty$,

A7. $\int_0^{+\infty} \int_0^{x} (xy)^{-1/2} f(x)g(y)dxdy < \infty$,

A8. $\int_0^{+\infty} x^{-1/2} f(x)G^2(x)dx < \infty$.

Theorem 3.1. *Let $\hat{R}_{GA}$ and $\hat{R}_{RIG}$ defined respectively by (8) and (9) be the SST-GA and SST-RIG estimators of the reliability parameter R with GA and RIG kernels, respectively. Under conditions **C1**, **C2** and A1-A7 given previously, the biases, variances and MSEs of the estimators are given by:*

(i) Biases

$$\begin{aligned} Bias\{\hat{R}_{GA}\} &= h_n \int_0^{+\infty} G(x)\{f'(x) + \frac{1}{2}xf''(x)\}dx + \frac{h_m}{2}\int_0^{+\infty} f(x)\{g(x) + xg'(x)\}dx \\ &+ \frac{h_n h_m}{2}\int_0^{+\infty} \{f'(x) + \frac{1}{2}xf''(x)\}\{g(x) + xg'(x)\}dx + o(h_n h_m). \end{aligned}$$

$$\begin{aligned} Bias\{\hat{R}_{RIG}\} &= \frac{h_n}{2}\int_0^{+\infty} xG(x)f''(x)dx + \frac{h_m}{2}\int_0^{+\infty} f(x)\{xg'(x) - g(x)\}dx \\ &+ \frac{h_n h_m}{4}\int_0^{+\infty} xf''(x)\{xg'(x) - g(x)\}dx + o(h_n h_m). \end{aligned}$$

(ii) Variances

$$\begin{aligned} Var\{\hat{R}_{GA}\} &= Var\{\hat{R}_{RIG}\} \\ &= \frac{1}{4\pi}(nm)^{-1}(h_n h_m)^{-1/2}\int_0^{+\infty}\int_0^{x} (xy)^{-1/2} f(x)g(y)dxdy \\ &+ o((nm)^{-1}(h_n h_m)^{-1/2}). \end{aligned}$$

(iii) Means Squared Errors (MSEs)

$$\begin{aligned} MSE\{\hat{R}_{GA}\} &= Bias^2\{\hat{R}_{GA}\} + Var\{\hat{R}_{GA}\} \\ &= \Big\{h_n \int_0^{+\infty} G(x)\{f'(x) + \frac{1}{2}xf''(x)\}dx \\ &+ \frac{h_m}{2}\int_0^{+\infty} f(x)\{g(x) + xg'(x)\}dx \\ &+ \frac{h_n h_m}{2}\int_0^{+\infty} \{f'(x) + \frac{1}{2}xf''(x)\}\{g(x) + xg'(x)\}dx\Big\}^2 \\ &+ \frac{1}{4\pi}(nm)^{-1}(h_n h_m)^{-1/2}\int_0^{+\infty}\int_0^{x} (xy)^{-1/2} f(x)g(y)dydx \\ &+ o((h_n h_m)^2 + (nm)^{-1}(h_n h_m)^{-1/2}) \end{aligned}$$

and

$$\begin{aligned} MSE\{\hat{R}_{RIG}\} &= \left\{\frac{h_n h_m}{4}\int_0^{+\infty} xf''(x)\{g(x)+xg'(x)\}dx + \frac{h_n}{2}\int_0^{+\infty} xG(x)f''(x)dx \right.\\ &+ \left. \frac{h_m}{2}\int_0^{+\infty} f(x)\{xg'(x)-g(x)\}dx\right\}^2 \\ &+ \frac{1}{4\pi nm}(h_n h_m)^{-1/2}\int_0^{+\infty}\int_0^{x}(xy)^{-1/2}f(x)g(y)dydx \\ &+ o((h_n h_m)^2 + (nm)^{-1}(h_n h_m)^{-1/2}). \end{aligned}$$

Proof. The proof is given in Appendix (Section A.1). □

3.3. Asymptotic properties

In this section, we establish the asymptotic properties of the proposed GA and RIG kernel estimators for SST under Lemma 3.2, the conditions **C1**, **C2** and the assumptions A1, A4, and A8.

Lemma 3.2. *If the PDF g follows condition **C1** and sequence $(h_m)_m$ follows condition **C2**. Under condition $\mathbb{E}[\int_0^x K_{k(y,h_m)}(Y)\,dy] < +\infty, \forall m, \forall x$, the following holds*

$$\hat{G}_k(x) \xrightarrow{a.s} G(x),$$

where $\hat{G}_k,\ k = GA$ or $k = RIG$ are the cdf kernel estimators of G and $\xrightarrow{a.s}$ denotes the almost sure convergence.

Proof. The proof is given in Appendix (Section A.2). □

Theorem 3.3. *Let $\hat{R}_{GA}$ and $\hat{R}_{RIG}$ defined respectively by (8) and (9) be the SST-GA and SST-RIG estimators of the reliability parameter R with GA and RIG kernels, respectively. The asymptotic biases, variances and MSEs of the estimators are given by:*

(i) Biases

$$Bias\{\hat{R}_{GA}\} = h_n\int_0^{+\infty} G(x)\{f'(x) + \frac{1}{2}xf''(x)\}dx + o(h_n),$$

$$Bias\{\hat{R}_{RIG}\} = \frac{h_n}{2}\int_0^{+\infty} xf''(x)G(x)dx + o(h_n).$$

(ii) Variances

$$Var\{\hat{R}_{GA}\} = Var\{\hat{R}_{RIG}\} = \frac{1}{2\sqrt{\pi}}n^{-1}h_n^{-1/2}\int_0^{+\infty} x^{-1/2}f(x)G^2(x)dx + o(n^{-1}h_n^{-1/2}).$$

(iii) Means Squared Errors (MSEs)

$$\begin{aligned} MSE\{\hat{R}_{GA}\} = h_n^2&\left\{\int_0^{+\infty} G(x)\{f'(x)+\frac{1}{2}xf''(x)\}dx\right\}^2 \\ &+ \frac{1}{2\sqrt{\pi}}n^{-1}h_n^{-1/2}\int_0^{+\infty} x^{-1/2}f(x)G^2(x)dx + o(h_n^2 + n^{-1}h_n^{-1/2}), \end{aligned} \tag{10}$$

and

$$\begin{aligned} MSE\{\hat{R}_{RIG}\} = \frac{h_n^2}{4}&\left\{\int_0^{+\infty} G(x)xf''(x)dx\right\}^2 + \frac{1}{2n\sqrt{\pi}}h_n^{-1/2}\int_0^{+\infty} x^{-1/2}f(x)G^2(x)dx \\ &+ o(h_n^2 + n^{-1}h_n^{-1/2}). \end{aligned} \tag{11}$$

Proof. The proof is given in Appendix (Section A.3). □

By minimizing the MSEs given in (10) and (11) with respect to h, we obtain the optimal bandwidths for the estimators $\hat{R}_{GA}$ and $\hat{R}_{RIG}$:

$$h_{GA}^{opt} = \frac{\left(\frac{1}{2\sqrt{\pi}}\int_0^{+\infty} x^{-1/2} f(x)G^2(x)dx\right)^{2/5}}{4^{2/5}\left(\int_0^{+\infty} G(x)\{f'(x)+\frac{1}{2}xf''(x)\}dx\right)^{4/5}} n^{-2/5} \tag{12}$$

and

$$h_{RIG}^{opt} = \left(\frac{\frac{1}{2\sqrt{\pi}}\int_0^{+\infty} x^{-1/2} f(x)G^2(x)dx}{\left\{\int_0^{+\infty} xf''(x)G(x)dx\right\}^2}\right)^{2/5} n^{-2/5}. \tag{13}$$

Substituting optimal bandwidths given in (12) and (13) in formulas (10) and (11), respectively, the optimal asymptotic means squared errors can be written as

$$MSE^*\{\hat{R}_{GA}\} = \frac{5}{4^{4/5}}\left[\int_0^{+\infty} G(x)\{f'(x)+\frac{1}{2}xf''(x)\}dx\right]^{2/5}\left[\frac{1}{2\sqrt{\pi}}\int_0^{+\infty}\frac{f(x)G^2(x)}{\sqrt{x}}dx\right]^{4/5} n^{-4/5}$$

and

$$MSE^*\{\hat{R}_{RIG}\} = \frac{5}{4}\left[\int_0^{+\infty} xf''(x)G(x)dx\right]^{2/5}\left[\frac{1}{2\sqrt{\pi}}\int_0^{+\infty}\frac{f(x)G(x)^2}{\sqrt{x}}dx\right]^{4/5} n^{-4/5}.$$

3.4. Bandwidth selection

Any kernel reliability estimator must have an appropriately selected bandwidth in order to be used. In this section, we present two different selection methods: rule-of-thumb (RT) and unbiased cross-validation (UCV).

Rule-of-thumb method

This method involves minimizing the MSEs given in (10) and (11) as a function of h. As the optimal bandwidths provided by (12) and (13) depend on the unknown quantities f, f', f" and G, they cannot be used in practice. To solve this problem, we suggest using Silverman's concept (Silverman 2018) of replacing the unknown PDF f and CDF G with a known lognormal reference model with parameters (μ_1, σ_1^2) and (μ_2, σ_2^2), respectively, such that

$$f_{(\mu_1,\sigma_1^2)}(x) = \frac{1}{x\sigma_1\sqrt{2\pi}}\exp\left(\frac{-(\ln x-\mu_1)^2}{2\sigma_1^2}\right), \quad x>0$$

and

$$G_{(\mu_2,\sigma_2^2)}(x) = \int_0^x f_{(\mu_2,\sigma_2^2)}(t)\,dt.$$

The first and second derivatives of f are respectively given by

$$f'_{(\mu_1,\sigma_1^2)}(x) = \frac{-1}{\sigma_1\sqrt{2\pi}}\left[\frac{1}{x^2}+\frac{\ln x-\mu_1}{x^2\sigma_1^2}\right]\exp\left(\frac{-(\ln x-\mu_1)^2}{2\sigma_1^2}\right), \quad x>0$$

and

$$f''_{(\mu_1,\sigma_1^2)}(x) = \frac{1}{x^3\sigma_1^3\sqrt{2\pi}}\left[3(\ln x-\mu_1)-1+2\sigma_1^2+\frac{(\ln x-\mu_1)^2}{2\sigma^2}\right]\exp\left(\frac{-(\ln x-\mu_1)^2}{2\sigma_1^2}\right), \quad x>0.$$

In practice, the parameters (μ_1, σ_1^2) and (μ_2, σ_2^2) can be replaced by their estimates $(\hat{\mu_1}, \hat{\sigma}_1^2)$ and $(\hat{\mu_2}, \hat{\sigma}_2^2)$, obtained from the maximum likelihood method.
Replacing (12) and (13), the optimal bandwidths (h_{RT}^{GA} and h_{RT}^{RIG}) are given by

$$h_{RT}^{GA} = \frac{\left(\frac{1}{2\sqrt{\pi}} \int_0^{+\infty} x^{-1/2} f_{(\hat{\mu}_1,\hat{\sigma}_1^2)}(x) G^2_{(\hat{\mu}_2,\hat{\sigma}_2^2)}(x)dx\right)^{2/5}}{4^{2/5}\left(\int_0^{+\infty} G_{(\hat{\mu}_2,\hat{\sigma}_2^2)}(x)\{f'_{(\hat{\mu}_1,\hat{\sigma}_1^2)}(x) + \frac{1}{2}xf''_{(\hat{\mu}_1,\hat{\sigma}_1^2)}(x)\}dx\right)^{4/5}} n^{-2/5}$$

$$h_{RT}^{RIG} = \left(\frac{\frac{1}{2\sqrt{\pi}} \int_0^{+\infty} x^{-1/2} f_{(\hat{\mu}_1,\hat{\sigma}_1^2)}(x) G^2_{(\hat{\mu}_2,\hat{\sigma}_2^2)}(x)dx}{\left\{\int_0^{+\infty} xf''_{(\hat{\mu}_1,\hat{\sigma}_1^2)}(x) G_{(\hat{\mu}_2,\hat{\sigma}_2^2)}(x)dx\right\}^2}\right)^{2/5} n^{-2/5}.$$

Other reference models can also be used, e.g. Gamma, Weibull, Birnbaum-Saunders, etc.

Unbiased cross validation method

This method consists in optimization of the squared error (SE) of the estimators $\hat{R}_k$, where k denote GA or RIG, that is given by

$$SE(\hat{R}_k) = (\hat{R}_k - R)^2 = \hat{R}_k^2 - 2\hat{R}_k R + R^2,$$

The last term of the expression does not depend on bandwidths (h_n, h_m), so we need to minimize the score functions CV_k, given by

$$CV_k(h_n, h_m) = \hat{R}_k(\hat{R}_k - 2\mathbb{E}[G(X)]).$$

We replace G by its estimators $\hat{G}_k$, then we get the new expressions of $CV_k(h)$, given by

$$UCV_k(h_n, h_m) = \frac{1}{nm}\sum_{i=1}^{n}\sum_{j=1}^{m}\int_0^{+\infty}\int_0^{x} K_{k(x,h_n)}(X_i)K_{k(y,h_m)}(Y_j)dydx \times$$
$$\left[\frac{1}{nm}\sum_{i=1}^{n}\sum_{j=1}^{m}\int_0^{+\infty}\int_0^{x} K_{k(x,h_n)}(X_i)K_{k(y,h_m)}(Y_j)dydx - \frac{2}{n(m-1)}\sum_{i=1}^{n}\sum_{j\neq l}^{m}\int_0^{x_i} K_{k(y_j,h_m)}(Y_l)dy_j\right].$$

The UCV optimal bandwidths h_{ucv}^{GA} and h_{ucv}^{RIG} are defined as

$$h_{ucv}^{k} = \arg\min_{h>0} UCV_k(h_n, h_m),$$

replacing k by GA and RIG, such that K_{GA} and K_{RIG} are the GA and RIG kernels defined in (2) and (4), respectively.

4. Simulation study

In this section, we investigate and compare the performance of the proposed SST estimators, based on the GA and RIG kernels developed in Section 3. The comparison concerns the GA and RIG kernels and the choice of smoothing parameter between the UCV and RT methods developed in Section 3.4.
We simulate data from non-negative lifetime distributions: Gamma (D1), Weibull (D2), Log-normal (D3) and Generalized Pareto (D4). The corresponding PDFs are listed in Table 1.
We note that the random variables X and Y are independent and do not necessarily have the same probability distribution. We then consider the following situations :

$$\begin{array}{lll} D1*D1: & X \sim D1(3,1), & Y \sim D1(1.5,1) \\ D2*D1: & X \sim D2(0.5,2), & Y \sim D1(1.5,1) \\ D3*D4: & X \sim D3(1,1), & Y \sim D4(0.5,0.9) \\ D4*D4: & X \sim D4(1.5,0.9), & Y \sim D4(0.5,0.9) \end{array}$$

Table 1: The probability distributions

	Distribution	Density
D1	Gamma (α, β)	$f(x) = \dfrac{x^{\alpha-1} e^{x/\beta}}{\gamma(\alpha)\beta^{\alpha}}$
D2	Weibull (λ, k)	$f(x) = \dfrac{k}{\lambda}\left(\dfrac{x}{\lambda}\right)^{k-1} \exp\left[-\left(\dfrac{x}{\lambda}\right)^{k}\right]$
D3	Lognormal (μ, σ^2)	$f(x) = \dfrac{1}{x\sigma\sqrt{2\pi}} \exp\left(\dfrac{-(\ln x - \mu)^2}{2\sigma^2}\right)$
D4	Generalized Pareto (α, λ)	$f(x) = \alpha\lambda(1 + \lambda x)^{-(\alpha+1)}$

The R reliability values calculated for the four previous models are , respectively $R_1 = 0.78, R_2 = 0.42, R_3 = 0.76, R_4 = 0.68$. For each model selected, 100 replications of equal or different sizes are generated: $(n, m) =$ (10, 10), (100, 100), (200, 200), (500, 500), (1000, 1000), (200, 10), (200, 100), (200, 500), (200, 1000), (10, 200), (100, 200), (500, 200), (1000, 200). The criterion used to compare the performance of the proposed estimators is the squared error (SE) criterion defined by

$$SE = (\hat{R} - R)^2.$$

Tables 2 and 3 present the average SE based on 100 replications for the SST reliability estimators for the models D1*D1, D2*D1, D3*D4 and D4*D4, using the bandwidths h_{UCV} and h_{RT} obtained by the UCV and RT methods, for different and equal sample sizes, respectively. All calculations were performed using R software. In terms of average SE, the obtained results based on GA and RIG kernels reveal globally that mean SE values decrease as sample sizes n and m increase. The best results are mainly obtained with the RIG kernel combined with the RT bandwidth selection method, compared with the UCV method. In fact:

- In the case $n = m$, most SE values decrease as the sample size increases. The SE values also decrease for $n \neq m$, except some cases where the strength size n is fixed while the stress size m varies.
- The RIG kernel gives the best results, whatever sample sizes, smoothing method and model used, except for the D4*D4 model combined with the RT method for medium and large sample sizes $(500, 500)$ and $(1000, 1000)$, and for $(1000, 200)$.
- The RT method gives around 60% of the best results in each of the two tables (for equal and different sizes). It should be noted that for equal, small or medium sizes (between 10 and 200), it is not possible to choose the best method; one of the two methods can then be used indifferently.
- The D2*D1 model has the lowest SE values, whatever the smoothing parameter, kernel and sample sizes used.

5. Application to real data

In this section, we use real data to test the proposed SST-GA and SST-RIG estimators, as well as the two bandwidth selection methods, RT and UCV. Here is a description of the data set considered:

Data set: The data set deals with the amount of annual rainfall (in inches) recorded at the Los Angeles Civic Center that are available in the web site of Los Angeles Almanac (http://www.laalmanac.com), on two 25 year periods. These data are discussed by Tarvirdizade and Ahmadpour (2016) in the context of parametric reliability estimation based

Table 2: Some expected values of (SE) for R estimators, based on 100 replications for considered models in simulation, using the bandwidths h_{UCV} and h_{RT}, for different sample sizes

		D1*D1		D2*D1		D3*D4		D4*D4	
(n, m)	kernel	RT	UCV	RT	UCV	RT	UCV	RT	UCV
(200,10)	GA	0.02510	0.12045	0.00795	0.02668	0.O3601	0.02319	0.01192	0.04927
	RIG	**0.01471**	**0.01425**	**0.00333**	**0.00291**	**0.00266**	**0.00500**	**0.00995**	**0.01095**
(200,100)	GA	0.01650	0.11572	0.00549	0.02758	0.01876	0.03313	0.00642	0.03102
	RIG	**0.01069**	**0.01006**	**0.00079**	**0.00101**	**0.00184**	**0.00225**	**0.00623**	**0.00453**
(200,500)	GA	0.01743	0.14782	0.00576	0.03913	0.03497	0.04946	0.00778	0.06741
	RIG	**0.01073**	**0.01915**	**0.00125**	**0.00122**	**0.00249**	**0.00374**	**0.00673**	**0.00756**
(200,1000)	GA	0.02196	0.15585	0.00829	0.04043	0.03681	0.05015	0.00937	0.06697
	RIG	**0.01100**	**0.02062**	**0.00130**	**0.00171**	**0.00321**	**0.00564**	**0.00761**	**0.00828**
(10,200)	GA	0.12255	0.16198	0.01530	0.06475	0.15476	0.05230	0.03903	0.08197
	RIG	**0.02587**	**0.02391**	**0.00663**	**0.00750**	**0.01274**	**0.00542**	**0.01231**	**0.01206**
(100,200)	GA	0.03922	0.15606	0.01118	0.03924	0.05911	0.04941	0.O1888	0.07843
	RIG	**0.01688**	**0.01672**	**0.00179**	**0.00487**	**0.00233**	**0.00420**	**0.00859**	**0.00967**
(500,200)	GA	0.01350	0.11045	0.00252	0.02721	0.03184	0.02091	**0.00385**	0.02852
	RIG	**0.00813**	**0.00616**	**0.00074**	**0.00246**	**0.00149**	**0.00175**	0.00413	**0.00544**
(1000,200)	GA	0.00918	0.08927	0.00080	0.01551	0.02167	0.01559	**0.00255**	0.00763
	RIG	**0.00583**	**0.00519**	**0.00061**	**0.00142**	**0.00083**	**0.00128**	0.00306	**0.00450**

Note: Bold values indicate the best results.

Table 3: Some expected values of (SE) for R estimators, based on 100 replications for considered models in simulation, using the bandwidths h_{UCV} and h_{RT}, for equal sample sizes

		D1*D1		D2*D1		D3*D4		D4*D4	
(n, m)	kernel	RT	UCV	RT	UCV	RT	UCV	RT	UCV
(10,10)	GA	0.18941	0.15553	0.03553	0.04067	0.13056	0.04133	0.07244	0.04947
	RIG	**0.02546**	**0.01253**	**0.00708**	**0.02042**	**0.01062**	**0.01084**	**0.01620**	**0.01095**
(100,100)	GA	0.06228	0.13408	0.00559	0.03263	0.04690	0.02709	0.01899	0.04628
	RIG	**0.01568**	**0.01194**	**0.00180**	**0.00355**	**0.00307**	**0.00496**	**0.00742**	**0.00620**
(200,200)	GA	0.01950	0.11213	0.00358	0.03004	0.02913	0.01949	0.01106	0.03796
	RIG	**0.01085**	**0.00832**	**0.00119**	**0.00382**	**0.00232**	**0.00240**	**0.00663**	**0.00531**
(500,500)	GA	0.00835	0.10664	0.00197	0.02767	0.01869	0.01492	**0.00276**	0.03598
	RIG	**0.00749**	**0.00749**	**0.00050**	**0.00074**	**0.00160**	**0.00147**	0.00375	**0.00400**
(1000,1000)	GA	0.00726	0.10288	0.00139	0.01222	0.01424	0.01011	**0.00143**	0.03039
	RIG	**0.00515**	**0.00384**	**0.00031**	**0.00058**	**0.00076**	**0.00090**	0.00266	**0.00304**

Note: Bold values indicate the best results.

on upper record values from a two-parameter bathtub-shaped lifetime distribution, then estimating the reliability R based on the inferred distribution. In contrast, the nonparametric approach proposed here does not assume a specific functional form for the stress and strength distributions. It uses kernel density estimators tailored for positive data (namely, GA and RIG kernels) to estimate R directly. The bandwidth parameter h was selected either by the rule of thumb method (RT) or using unbiased cross-validation (UCV). The data values are:

Data Set 1 (from 1959 to 1983): 8.18, 4.85, 18.79, 8.38, 7.93, 13.68, 20.44, 22.00, 16.58, 27.47, 7.74, 12.32, 7.17, 21.26, 14.92, 14.35, 7.21, 12.30, 33.44, 19.67, 26.98, 8.96, 10.71, 31.28, 10.43.

Data Set 2 (from 1984 to 2008): 12.82, 17.86, 7.66, 2.48, 8.08, 7.35, 11.99, 21.00, 7.36, 8.11, 24.35, 12.44, 12.40, 31.01, 9.09, 11.57, 17.94, 4.42, 16.42, 9.25, 37.96, 13.19, 3.21, 13.53, 9.08.

In the context of our analysis, stress and strength are conceptual terms often used in reliability engineering and applied to rainfall data for comparative analysis. Here, strength represents the basic or historical "capacity" of the system. It reflects the behavior of rainfall regimes in the past and serves as a reference for comparison with "stress" data. It is therefore used to assess the resilience or stability of rainfall regimes in response to stress. If recent precipitation (stress) exceeds historical trends (strength), this may suggest increased variability or extreme events in the recent period. Consequently, strength represents the historical baseline (1959-1983), and stress represents the recent observations (1984-2008), allowing for a comparative analysis of rainfall trends over time.

Firstly, we applied the Mann-Whitney (MW) test on the strength and stress data in order to test the independence between the variables X and Y, respectively. The value of the MW statistic equals 367 with the corresponding p-value of 0.2948. This indicates that the stress and force variables are independent. We also present, in Table 4, the descriptive statistics linked to these data. We can see that the data is positively skewed and exhibits a medium level of Kurtosis. Therefore, we can apply the method proposed in this paper for SST system reliability estimation using GA and RIG kernels.
Table 5 presents the estimated reliability values $\hat{R}_{RT}$ and $\hat{R}_{UCV}$ (fourth and sixth columns) by the SST-GA and SST-RIG estimators, as well as the bandwidth selectors h_{RT} and h_{UCV} (third and fifth columns), obtained by the RT and UCV methods, respectively, for the real data set. Recall that the RT method is based on the lognormal reference model (see Section 3.4). We find that the model reliability estimate is on average 0.5795, with a slight overestimation observed for the Gamma kernel with h_{RT}. While the parametric analysis yields a reliability value of around 0.43, for the different methods used (maximum likelihood and the Bayesian approaches), the non-parametric estimation falls within the range of 0.5786 to 0.5862. This difference reflects the methodological nature of non-parametric techniques, which rely solely on observed data without assuming a specific underlying distribution. This approach can capture features that parametric models, limited by their assumptions, might overlook. Therefore, the non-parametric estimate serves as a valuable complement to the parametric result, enriching the analysis and offering a broader view of the system's reliability.

Table 4: Descriptive statistic for the indicated data set. CV: coefficient of variation, CS: coefficient of skewness and CK: coefficient of kurtosis

Data set	Mean	Median	SD	CV	CS	CK	Min	Max	$n = m$
Strength	15.481	13.68	8.075	0.521	0.735	2.510	4.85	33.44	25
Stress	13.222	11.99	8.336	0.630	1.384	4.758	2.48	37.96	25

Table 5: Estimation of SST reliability model with real data

Data set	kernel	h_{RT}	$\hat{R}_{RT}$	h_{UCV}	$\hat{R}_{UCV}$
	GA	0.1892543	0.5861947	1.39735	0.5786196
	RIG	0.8207905	0.5813191	1.098154	0.57861191

Reliability R being a parameter, we cannot give a graphical representation for its nonparametric estimation, which is based on kernel estimation of the PDFs of the stress and strength variables. Figure 1 shows the PDFs estimates for real data based on The GA and RIG kernels combined with RT and UCV bandwidth selectors. The two kernel estimates of the PDFs f and g are plotted on the histograms of the corresponding data sets, the histogram being the actual estimate of the data distribution.

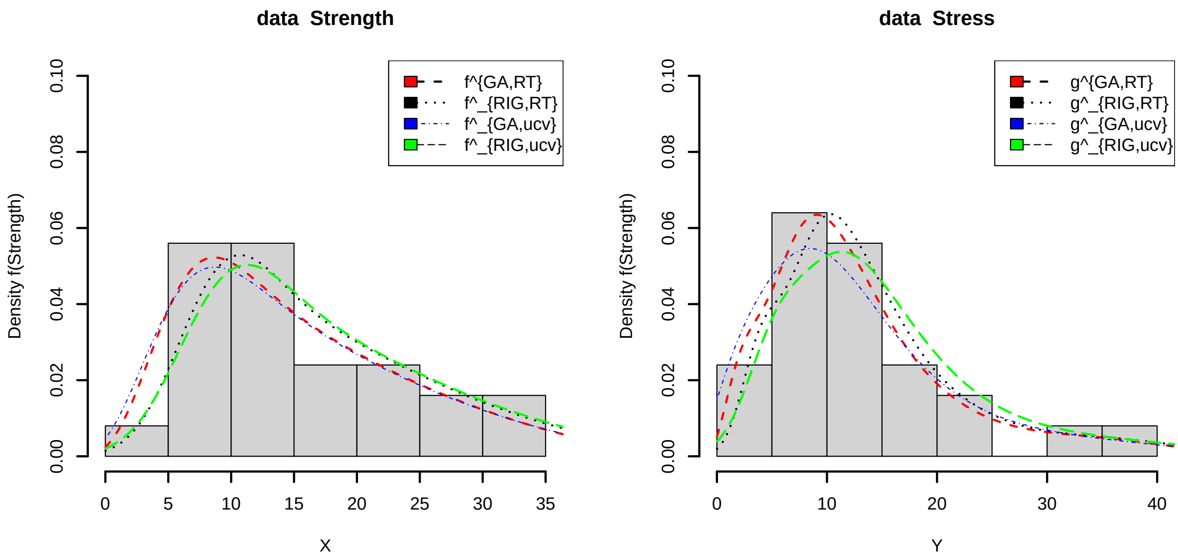

Figure 1: PDF estimation of real data by GA and RIG kernels using RT and UCV methods

As shown in Figure 1, the GA and RIG kernel estimates give a good account of the shape of the histograms, for all bandwidth selectors. We can observe that, in terms of smoothing quality, the best performance is obtained with the RT method for the RIG and GA kernel, smoothing quality is very satisfactory for the UCV and RT methods with both kernels. Taking into account the estimation adjustments considered, we can say that the proposed kernels provide good PDFs estimation when combined with the RT method.
The graphical and simulation results allow us to conclude that the RIG kernel combined with the RT method gives a good estimate of the probability densities of the stress and strength variables, as well as model reliability.

6. Conclusion

This paper proposes an alternative to parametric approach for estimating the stress strength reliability model. We consider the nonparametric kernel method based on Gamma and Reciprocal Inverse Gaussian kernels, introduced by Chen (2000) and Scaillet (2004), respectively. The proposed SST estimators have good statistical and asymptotic properties, and their performances are investigated through simulation study for independent non-negative data, which is conducted for different lifetime distributions (Gamma, Weibull, Lognormal, and Generalized Pareto), sample sizes and bandwidth selection methods (RT and UCV). In general, the SST-RIG estimator outperforms the SST-GA estimator in the sense of SE, and the rule of the thumb approach outperforms the unbiased cross validation approach for bandwidth selector. The kernel based nonparametric estimators offer a robust and accurate alternative to parametric methods in SST reliability estimation. Their flexibility and ease of implementation make them valuable tools, particularly when distributional assumptions are uncertain. Finally, other asymmetric kernels combined with other smoothing parameter selection methods can be investigated. This work could constitute a basis for a vast field of research in nonparametric estimation applied to SST models.

References

Abu El Azm WS, Almetwally EM, Alghamdi AS, Aljohani HM, Muse AH, Abo-Kasem OE (2021). "Stress–strength Reliability for Exponentiated Inverted Weibull Distribution with Application on Breaking of Jute Fiber and Carbon Fibers." *Computational Intelligence and Neuroscience*, **2021**(1), 4227346. doi:10.1155/2021/4227346.

Almarashi AM, Algarni A, Nassar M (2020). "On Estimation Procedures of Stress–strength Reliability for Weibull Distribution with Application." *PLOS One*, **15**(8), e0237997. doi:10.1371/journal.pone.0237997.

Asgharzadeh A, Valiollahi R, Raqab MZ (2013). "Estimation of the Stress–strength Reliability for the Generalized Logistic Distribution." *Statistical Methodology*, **15**, 73–94. doi:10.1016/j.stamet.2013.05.002.

Bouezmarni T, El Ghouch A, Mesfioui M (2011). "Gamma Kernel Estimators for Density and Hazard Rate of Right-Censored Data." *Journal of Probability and Statistics*, **2011**(1), 937574. doi:10.1155/2011/937574.

Bouezmarni T, Scaillet O (2005). "Consistency of Asymmetric Kernel Density Estimators and Smoothed Histograms with Application to Income Data." *Econometric Theory*, **21**(2), 390–412. doi:10.1017/S0266466605050218.

Chekkal S, Lagha K, Zougab N (2023). "Generalized Birnbaum–Saunders Kernel for Hazard Rate Function Estimation." *Communications in Statistics-Simulation and Computation*, **52**(4), 1546–1561. doi:10.1080/03610918.2021.1887228.

Chen SX (1999). "Beta Kernel Estimators for Density Functions." *Computational Statistics & Data Analysis*, **31**(2), 131–145. doi:10.1016/S0167-9473(99)00010-9.

Chen SX (2000). "Probability Density Function Estimation Using Gamma Kernels." *Annals of the Institute of Statistical Mathematics*, **52**, 471–480. doi:10.1023/A:1004165218295.

Gustafsson J, Hagmann M, Nielsen JP, Scaillet O (2009). "Local Transformation Kernel Density Estimation of Loss Distributions." *Journal of Business & Economic Statistics*, **27**(2), 161–175. doi:10.1198/jbes.2009.0011.

Hagmann M, Scaillet O (2007). "Local Multiplicative Bias Correction for Asymmetric Kernel Density Estimators." *Journal of Econometrics*, **141**(1), 213–249. doi:10.1016/j.jeconom.2007.01.018.

Jia X, Nadarajah S, Guo B (2017). "Bayes Estimation of P(Y<X) for the Weibull Distribution with Arbitrary Parameters." *Applied Mathematical Modelling*, **47**, 249–259. doi:10.1016/j.apm.2017.03.020.

Jin X, Kawczak J (2003). "Birnbaum-Saunders and Lognormal Kernel Estimators for Modelling Durations in High Frequency Financial Data." *Annals of Economics and Finance*, **4**, 103–124.

Kotz S, Lumelskii Y, Pensky M (2003). *The Stress–strength Model and its Generalizations: Theory and Applications*. World Scientific.

Kundu D, Gupta RD (2006). "Estimation of P[Y< X] for Weibull Distributions." *IEEE Transactions on Reliability*, **55**(2), 270–280. doi:10.1109/TR.2006.874918.

Marchant C, Bertin K, Leiva V, Saulo H (2013). "Generalized Birnbaum–Saunders Kernel Density Estimators and an Analysis of Financial Data." *Computational Statistics & Data Analysis*, **63**, 1–15. doi:10.1016/j.csda.2013.01.013.

Muhammad I, Wang X, Li C, Yan M, Chang M (2020). "Estimation of the Reliability of a Stress–strength System from Poisson Half Logistic Distribution." *Entropy*, **22**(11), 1307. doi:10.3390/e22111307.

Parzen E (1962). "On Estimation of a Probability Density Function and Mode." *The Annals of Mathematical Statistics*, **33**(3), 1065–1076.

Raqab MZ, Madi MT, Kundu D (2008). "Estimation of P (Y<X) for the Three-parameter Generalized Exponential Distribution." *Communications in Statistics–Theory and Methods*, **37**(18), 2854–2864. doi:10.1080/03610920802162664.

Rezaei S, Tahmasbi R, Mahmoodi M (2010). "Estimation of P [Y<X] for Generalized Pareto Distribution." *Journal of Statistical Planning and Inference*, **140**(2), 480–494. doi:10.1016/j.jspi.2009.07.024.

Salha R (2013). "Estimating the Density and Fazard Rate Functions Using the Reciprocal Inverse Gaussian Kernel." In *15th Applied Stochastic Models and Data Analysis International Conference, Mataro (Barcelona) Spain.*

Scaillet O (2004). "Density Estimation Using Inverse and Reciprocal Inverse Gaussian Kernels." *Nonparametric Statistics*, **16**(1-2), 217–226. doi:10.1080/10485250310001624819.

Silverman BW (2018). *Density Estimation for Statistics and Data Analysis.* Routledge.

Tarvirdizade B, Ahmadpour M (2016). "Estimation of the Stress–strength Reliability for the Two–parameter Bathtub–shaped Lifetime Distribution Based on Upper Record Values." *Statistical Methodology*, **31**, 58–72. doi:10.1016/j.stamet.2016.01.005.

A. Appendix

We present a sketch of proofs of Theorem 3.1, Lemma 3.2 and Theorem 3.3.

A.1. Sketch of the proof of Theorem 3.1

(i) Biases

First, by using independence between X and Y we can write,

$$\begin{aligned}
\text{Bias}[\hat{R}_k] &= \mathbb{E}[\hat{R}_k] - R = \int_0^{+\infty}\int_0^x \mathbb{E}\{\hat{f}(x)\}\mathbb{E}\{\hat{g}(y)\}dydx - R \\
&= \int_0^{+\infty}\int_0^x \{\text{Bias}(\hat{f}(x)+f(x))\}\{\text{Bias}(\hat{g}(y)+g(y))\}dydx - R \\
&= \int_0^{+\infty}\int_0^x \text{Bias}\{\hat{f}(x)\}\text{Bias}\{\hat{g}(y)\}dydx + \int_0^{+\infty} G(x)\text{Bias}\{\hat{f}(x)\}dx \\
&\quad + \int_0^{+\infty}\int_0^x f(x)\text{Bias}\{\hat{g}(y)\}dydx \\
&= A_1 + A_2 + A_3.
\end{aligned}$$

where,

$$A_1 = \int_0^{+\infty}\int_0^x \text{Bias}\{\hat{f}(x)\}\text{Bias}\{\hat{g}(y)\}dydx, \quad A_2 = \int_0^{+\infty} G(x)\text{Bias}\{\hat{f}(x)\}dx \quad \text{and,}$$

$$A_3 = \int_0^{+\infty}\int_0^x f(x)\text{Bias}\{\hat{g}(y)\}dydx.$$

For Bias of the $\hat{R}_{GA}$ estimator

From relation (5) and using integration by part, we obtain

$$\begin{aligned}
A_1 &= \int_0^{+\infty}\int_0^x \left(h_n\left\{f'(x)+\frac{1}{2}xf''(x)\right\}+o(h_n)\right)\left(h_m\left\{g'(y)+\frac{1}{2}yg''(y)\right\}+o(h_m)\right)dydx \\
&= \int_0^{+\infty}\left(h_n\left\{f'(x)+\frac{1}{2}xf''(x)\right\}+o(h_n)\right)\left(\frac{h_m}{2}\{g(x)+xg'(x)\}+o(h_m)\right)dx \\
&= \frac{h_nh_m}{2}\int_0^{+\infty}\left\{f'(x)+\frac{1}{2}xf''(x)\right\}\{g(x)+xg'(x)\}dx + o(h_n+h_m+h_nh_m),
\end{aligned}$$

$$A_2 = \int_0^{+\infty} h_n\left\{f'(x)+\frac{1}{2}xf''(x)\right\}G(x)dx + o(h_n)$$

and,

$$A_3 = \frac{h_m}{2}\int_0^{+\infty} f(x)\{g(x)+xg'(x)\}dx + o(h_m).$$

For Bias of the $\hat{R}_{RIG}$ estimator

From relation (6) and using integration by part, we obtain

$$\begin{aligned}
A_1 &= \int_0^{+\infty}\int_0^x \left\{\frac{h_n}{2}xf''(x)+o(h_n)\right\}\left\{\frac{h_m}{2}yg''(y)+o(h_m)\right\}dydx \\
&= \frac{h_nh_m}{4}\int_0^{+\infty} xf''(x)\{xg'(x)-g(x)\}dx + o(h_n+h_m+h_nh_m)
\end{aligned}$$

$$A_2 = \frac{h_n}{2}\int_0^{+\infty} xf''(x)G(x)dx + o(h_n)$$

and,

$$A_3 = \frac{h_m}{2}\int_0^{+\infty} f(x)\left\{xg'(x) - g(x)\right\}dx + o(h_m).$$

(ii) variances
We have,

$$\begin{aligned}
\mathrm{Var}(\hat{R}_k) &= \mathrm{Var}\left(\frac{1}{nm}\sum_{i=1}^{n}\sum_{j=1}^{m}\int_0^{+\infty}\int_0^{x} K_{k(x,h_n)}(X_i)K_{k(y,h_m)}(Y_j)dydx\right)\\
&= \frac{1}{nm}\mathrm{Var}\left(\int_0^{+\infty}\int_0^{x} K_{k(x,h_n)}(X_1)K_{k(y,h_m)}(Y_1)dydx\right)\\
&= \frac{1}{nm}\int_0^{+\infty}\int_0^{x}\mathrm{Var}\left(K_{k(x,h_n)}(X_1)K_{k(y,h_m)}(Y_1)dydx\right)\\
&= \frac{1}{nm}\int_0^{+\infty}\int_0^{x}\mathbb{E}\left(K^2_{k(x,h_n)}(X_1)K^2_{k(y,h_m)}(Y_1)dydx\right)dydx\\
&- \frac{1}{nm}\int_0^{+\infty}\int_0^{x}\mathbb{E}^2(K_{k(x,h_n)}(X_1))\mathbb{E}^2(K_{k(y,h_m)}(Y_1))dydx,
\end{aligned}$$

where,

$$\mathbb{E}(K_{k(x,h_n)}(X_1)) = f(x) \quad \text{and} \quad \mathbb{E}(K_{k(y,h_m)}(Y_1)) = g(y).$$

Consequently,

$$\begin{aligned}
\mathrm{Var}(\hat{R}_k) &= \frac{1}{nm}\int_0^{+\infty}\int_0^{x}\mathbb{E}\left(K^2_{k(x,h_n)}(X_1)K^2_{k(y,h_m)}(Y_1)dydx\right) - \frac{1}{nm}\int_0^{+\infty}\int_0^{x}(f(x)g(y))^2dydx\\
&\simeq \frac{1}{nm}\int_0^{+\infty}\int_0^{x}\mathbb{E}[K^2_{k(x,h_n)}(X_1)]\mathbb{E}[K^2_{k(y,h_m)}(Y_1)]dydx - O\left(\frac{1}{nm}\right)\\
&= \int_0^{+\infty}\int_0^{x}\mathrm{var}\{\hat{f}(x)\}\mathrm{var}\{\hat{g}(y)\}dydx
\end{aligned}$$

and we have, $\mathrm{var}\{\hat{f}_{GA}(x)\} = \mathrm{var}\{\hat{f}_{RIG}(x)\}$, see Scaillet (2004).
This leads to the variance result.

A.2. Sketch of the proof of Lemma 3.2

First, we have

$$\begin{aligned}
\mathbb{E}[\hat{G}_k(x)] &= \mathbb{E}\left[\int_0^{x} K_{(y,h_m)}(Y)dy\right]\\
&= \int_0^{x}\int_0^{\infty} K_{k(y,h_m)}(u)g(u)dudy\\
&= \int_0^{x}\mathbb{E}[g(\xi_k)]dy, \qquad (14)
\end{aligned}$$

and ξ_k is the random variable follows a $GA(y, h_m)$ or $RIG(y, h_m)$ distributions, and following Chen (2000) and Scaillet (2004) we have,

$$\mathbb{E}[g(\xi_{GA})] = g(y) + h_m\{g'(y) + \frac{1}{2}yg''(y)\} + o(h_m)$$

and

$$\mathbb{E}[g(\xi_{RIG})] = g(y) + h_m\frac{1}{2}yg''(y) + o(h_m).$$

By replacing in (14), we have

$$\begin{aligned}
\mathbb{E}[\hat{G}_{GA}(x)] &= \int_0^x \{g(y) + h_m[g'(y) + \frac{1}{2}yg''(y)] + o(h_m)\}dy \\
&= \int_0^x g(y)dy + h_m \int_o^x \{g'(y) + \frac{1}{2}yg''(y)\}dy + o(h_m) \\
&= G(x) + \frac{h_m}{2}\{g(x) + xg'(x)\} + o(h_m) \\
&= G(x) + O(h_m). \qquad (15)
\end{aligned}$$

and

$$\begin{aligned}
\mathbb{E}[\hat{G}_{RIG}(x)] &= \int_0^x \{g(y) + h_m\frac{1}{2}yg''(y) + o(h_m)\}dy \\
&= G(x) + \frac{h_m}{2}\{xg'(x) - g(x)\} + o(h_m) \\
&= G(x) + O(h_m). \qquad (16)
\end{aligned}$$

Second, note that the m random variables $\int_0^x K_{K(y,h_m)}(Y_j)dx$ are i.i.d. as $\mathbb{E}[\int_0^x K_{K(y,h_m)}(Y_j)dx] < \infty, \forall m$. Hence, by the strong law of large numbers, we obtain,

$$\hat{G}_{GA}(x) - \mathbb{E}\left[\int_0^x K_{GA(y,h_m)}(Y_1)dy\right] \xrightarrow{a.s} 0, \qquad m \longrightarrow \infty, \qquad (17)$$

and

$$\hat{G}_{RIG}(x) - \mathbb{E}\left[\int_0^x K_{RIG(y,h_m)}(Y_1)dy\right] \xrightarrow{a.s} 0, \qquad m \longrightarrow \infty. \qquad (18)$$

Then, by using (15), (16), (17), (18) and the following classical decomposition:

$$\hat{G}_{GA}(x) - G(x) = [\mathbb{E}(\hat{G}_{GA}(x)) - G(x)] + [\hat{G}_{GA}(x) - \mathbb{E}(\hat{G}_{GA}(x))],$$

and

$$\hat{G}_{RIG}(x) - G(x) = [\mathbb{E}(\hat{G}_{RIG}(x)) - G(x)] + [\hat{G}_{RIG}(x) - \mathbb{E}(\hat{G}_{RIG}(x))],$$

we complete the proof of the lemma.

A.3. Sketch of the proof of Theorem 3.3

From the previous Lemma 3.2, we can write for m enough large that, for $k = GA$ or RIG

$$\hat{R}_k = \int_0^{+\infty} \hat{f}_k(x)\hat{G}_k(x)dx = \int_0^{+\infty} \hat{f}_k(x)G(x)dx \quad a.s.$$

Then for n enough large, the mean of $\hat{R}_k$ is simply given by

$$\mathbb{E}[\hat{R}_k] = \int_0^{+\infty} \mathbb{E}[\hat{f}_k(x)]G(x)dx.$$

Hence, the asymptotic bias is expressed as

$$\begin{aligned}
\text{Bias}[\hat{R}_k] &= \mathbb{E}[\hat{R}_k] - R \\
&= \int_0^{+\infty} \mathbb{E}[\hat{f}_k(x)]G(x)dx - \int_0^{+\infty} f(x)G(x)dx \\
&= \int_0^{+\infty} \text{bias}\{\hat{f}_k(x)\}G(x)dx.
\end{aligned}$$

Similarly, the asymptotic variance is given by

$$\text{Var}(\hat{R}_k) = \text{Var}\left[\int_0^{+\infty} \hat{f}_k(x)G(x)dx\right] = \int_0^{+\infty} G^2(x)\text{Var}\{\hat{f}_k(x)\}dx.$$

The results of theorem are obtained by replacing the expressions of the biases and variances of $\hat{f}_{GA}(x)$ and $\hat{f}_{RIG}(x)$ given in the Formulas (5),(6) and (7), respectively.

Affiliation:

Meriem Medjider
University of Bejaia, Faculty of Exact Sciences
Applied Mathematics Laboratory
06000 Bejaia, Algeria
E-mail: meriem.medjider@univ-bejaia.dz

Karima Lagha
University of Bejaia, Faculty of Exact Sciences
Research Unit LaMOS
06000 Bejaia, Algeria
E-mail: karima.lagha@univ-bejaia.dz

Austrian Journal of Statistics http://www.ajs.or.at/
published by the Austrian Society of Statistics http://www.osg.or.at/
Volume 55 *Submitted:* 2025-01-17
2026 *Accepted:* 2025-05-28