Journal of Statistics & Management Systems
ISSN 0972-0510 (Print), ISSN 2169-0014 (Online)
Vol. 29 (2026), No. 6, pp. 589–602
DOI : 10.47974/JSMS-1477

Estimating the reliability of a series stress-strength system using lognormal kernel

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

Karima Lagha †
Laboratory of Modeling and Optimization of Systems
Faculty of Exact Sciences
University of Bejaia
Bejaia 06000
Algeria

Abstract

In this paper, we have estimated the reliability of a system composed of elements connected in series and subject to the same stress, by the nonparametric kernel estimation method using lognormal kernel. This kernel is chosen for its advantageous properties, including an optimal convergence rate for the mean integrated squared error, non-negativity, free from boundary bias, and naturally varying in shape. Asymptotic properties such as bias, variance, and mean squared error have been established for the proposed estimator. Furthermore, we address the selection of the optimal bandwidth parameter an essential aspect of kernel estimation using both the rule of thumb and the unbiased cross validation techniques. Finally, a simulation study is carried out to highlight the performance of the system reliability estimator, based on lognormal kernel and to compare the two bandwidth selection techniques to determine the best one.

Subject Classification: *90B25, 30C40, 62G05.*

Keywords: *Stress strength model, Reliability, Lognormal kernel, Bandwidth parameter, Nonparametric estimation.*

* *ORCID-ID:* **0009-0006-6154-1259**
† *ORCID-ID:* **0000-0002-0274-9424**

* *E-mail:* **meriem.medjider@univ-bejaia.dz (Corresponding Author)**
† *E-mail:* **karima.lagha@univ-bejaia.dz**

1. Introduction

Nonparametric statistics is a branch of statistics that does not assume a particular shape for the data distribution. It is often used for hypothesis testing, group comparisons, or exploratory analyses when the conditions for parametric methods are not met. Recent advancements in nonparametric statistical methods have highlighted their practical advantages in reliability analysis and industrial applications. For instance, Aslam [1] proposed a nonparametric sign sampling plan for product acceptance determinations, offering a flexible and assumption-free framework suitable for uncertain production environments. Likewise, Ozen et al. [2] investigated the use of bootstrap techniques for estimating the variance of the sample median, emphasizing the robustness of nonparametric approaches in situations where classical methods may falter.

Nonparametric estimation is a component of nonparametric statistics. It focuses on building models or representations of data. It allows functions to be modelled without making strong parametric assumptions. Kernel-based nonparametric estimation is a flexible method for estimating probability density functions without relying on strict distributional assumptions. It belongs to a broader class of nonparametric techniques that are particularly useful when the underlying data structure is complex or unknown (see Chen [3, 4], Scaillet [5], Jin and Kawczak [6] and Marchant et al. [7]).This method has been widely applied across various fields, especially in reliability analysis, where it is used to estimate reliability and failure rate functions (see Bouezmarni, El Ghouch, and Mesfioui [8], Salha [9], and Chekkal, Lagha, and Zougab [10]).

Motivated by this growing interest in nonparametric inference, we consider in this paper, the problem of estimating the reliability of the series stress-strength system. The stress-strength model has been widely studied in the statistical literature. In a stress-strength system composed of a single component, the system reliability is defined as the probability $R = P(X > Y)$ where X denotes the strength and Y the applied stress. X and Y are random variables assumed to be independent with probability densities functions (PDFs) f and g, respectively. The reliability R of the stress-strength component and is a performance measure of the component. This model highlights the stress-strength relationship, with failure occurring when stress exceeds strength, and can be expressed as

$$R = P(X > Y) = \int_0^{+\infty} \int_0^{x} f(x)g(y)\,dydx = \int_0^{+\infty} f(x)G(x)\,dydx. \quad (1)$$

where, $G(x) = \int_0^{x} g(y)dy$ is the cumulative distribution function (CDF) of Y calculated at x.

Stress-strength reliability analysis has been extensively studied using parametric models with various life distributions. Notable examples include the works of Kundu and Gupta [11, 12], Mokhlis [13], Raqab, Madi and Kundu [14], Rezaei, Tahmasbi and Mahmoodi [15], Asgharzadeh, Valiollahi and Raqab [16], Rao et al. [17], Jia, Nadarajah and Guo [18], Muhammad et al.[19], Almarashi, Algarni, and Nassar [20] and Temraz [21], and others . In these studies, the random variables X and Y are presumed to be members of the same family of life distributions, but with distinct parameters. This assumption allows the application of parametric estimation methods like the Bayesian approach and maximum likelihood estimation.

Consider a series stress-strength system model, composed of s stochastically independent components placed in series and subjected to a stress Y. The resulting system functions correctly if and only if the strength of each component among the s exceeds the stress applied to the system. We note X_l the strength of the component l, for $l = 1, \dots, s$. The probability that the system does not fail is then given by

$$R = P(X > Y),$$

where $X = \min_{1 \le l \le s} X_l$, represents the strength of the system.

Since the strength and stress variables $X_l, l = \overline{1, s}$ and Y are non-negative, it is appropriate to use asymmetric kernels with non-negative support instead of symmetric kernels when estimating the reliability of a series stress-strength system. This choice helps to avoid the boundary effect, which refers to the high bias typically observed near the boundaries. The estimation of series stress-strength system reliability is carried out using the lognormal (LogN) kernel proposed by Jin and Kawczak [6], through the nonparametric estimation of the probability density functions of the variables $X_l, l = \overline{1, s}$ and Y. The LogN kernel estimator achieves the optimal rate of convergence for the mean integrated squared error, as it is boundary bias-free, non-negative, and inherently adaptive in shape, see Jin and Kawczak [6].

The structure of this paper is as follows: Section 2 provides a brief overview of the LogN kernel estimator for a probability density function f. Section 3 introduces the nonparametric LogN kernel estimator for the reliability of series stress-strength systems and examines its asymptotic properties, including bias, variance, and mean square error, along with bandwidth selection using the unbiased cross-validation and rule of thumb methods. Section 4 reports Monte Carlo simulation results. Section 5 concludes.

2. Preliminaries

Let Z be a continuous and positive random variable (r.v.) with an unknown PDF f defined on the interval $[0, \infty)$. Let $Z_1, \dots, Z_n$ be a random sample on Z, continuous, independent and identically distributed (i.i.d). The nonparametric LogN kernel estimator of the PDF is given by Jin and Kawczak [6]:

$$\hat{f}_{LogN}(z) = \frac{1}{n} \sum_{i=1}^{n} K_{LN(\ln z, 4\ln(1+h))}(Z_i), \tag{2}$$

where $h > 0$ is a bandwidth (or smoothing parameter) and $K_{LN(\ln z, 4\ln(1+h))}(.)$ is the LogN kernel defined by

$$K_{LogN(\ln z, 4\ln(1+h))}(t) = \frac{1}{\sqrt{8\pi \ln(1+h)}\, t} exp\left[-\frac{(\ln t - \ln z)^2}{8\ln(1+h)}\right], t > 0. \tag{3}$$

The expressions of the bias and variance for $\hat{f}_{LogN}$ are respectively derived by Jin and Kawczak [6], under the following conditions:

C1. The function f that is twice differentiable, with its second derivative both continuous and bounded,

C2. The bandwidth $h = h_n$ satisfies $\lim_{n\to\infty} h = 0$ and $\lim_{n\to\infty} nh = +\infty$.

$$Bias(\hat{f}_{LogN}(z)) = h[2zf'(z) + 2z^2 f''(z)] + o(h) \tag{4}$$

$$Var(\hat{f}_{LogN}(z)) = \frac{1}{4\sqrt{\pi}} h^{-1/2} n^{-1} z^{-1} f(z) + o(n^{-1}h^{-1/2}). \tag{5}$$

3. Serial stress-strength system kernel estimator

In this section, we develop the reliability estimation for the series stress-strength system using the LogN kernel. We discuss several properties of the proposed estimator. Additionally, we study bandwidth selection for the series stress-strength system reliability estimator using the popular rule of thumb and the unbiased cross-validation approach.

3.1 *Determining the Estimator*

The series stress-strength system can only function properly if all its components withstand the applied stress, due to its series structure. If any component fails, the entire system ceases to operate. In this context, the system strength X is a random variable such that

$$X = \min_{l=1}^{s} X_l,$$

where $X_l, l = \overline{1, s}$ are strength variables associated to the s components of the series stress-strength system subject to the stress Y.

Assuming that the components operate independently, the series stress-strength system's reliability is then:

$$R_s = P(X > Y) = P(\min_{l=1}^{s} X_l > Y) = P(\forall l, X_l > Y) = \prod_{l=1}^{s} R_l,$$

where $R_l = P(X_l > Y)$ and X_l is is the strength variable with PDF f_l assumed to be independent of random stress Y with CDF G for all $l = \overline{1, s}$, and $X_1, X_2, \ldots, X_s$ are independent random varaibles.

By substituting the expression of R_l given in (1) we can write,

$$R_s = \prod_{l=1}^{s} \int_0^{+\infty} \int_0^{x} f_l(x) g(y) dy dx = \prod_{l=1}^{s} \int_0^{+\infty} f_l(x) G(x) dx.$$

Let $X_{1l}, X_{2l}, \ldots, X_{nl}$ be a random sample on $X_l, l = \overline{1, s}$ from unknown PDF f_l, assumed to be independent of the random sample on Y given by $Y_1, Y_2, \ldots, Y_m$ from unknown PDF g and unknown CDF G, where $nl, l = \overline{1, s}$ and m are the respective sizes of the samples. The unknown PDFs $f_l, l = \overline{1, s}$ and g can be estimated by the LogN kernel estimators given in (2). In addition, the estimator of the unknown CDF G is given by

$$\hat{G}_{LogN}(x) = \int_0^x \hat{g}(y) dy = \int_0^x \frac{1}{m} \sum_{j=1}^{m} K_{LogN(\ln y, 4\ln(1+h_m))}(Y_j) dy.$$

The series stress-strength system reliability using the LogN kernel is given by

$$\hat{R}_s = \prod_{l=1}^{s} \frac{1}{nm} \int_0^{+\infty} \int_0^{x} \sum_{i=1}^{nl} \sum_{j=1}^{m} K_{LogN(\ln y, 4\ln(1+h_m))}(Y_j) K_{LogN(\ln x, 4\ln(1+h_n))}(X_{il}) dydx, \quad (6)$$

We assume the same size n for all samples of the strengths, then $h_n = h_{nl}, l = \overline{1,s}$

3.2 *Main results*

The theorem 3.2 below provides the asymptotic bias, variance and mean square error (MSE) of the estimator $\hat{R}_s$, defined in formula (6) under conditions **C1**, **C2**, the proposition 3.1 and additional assumptions stated below

A1. $\int_0^{+\infty} xG(x)f_l'(x)dx < \infty$

A2. $\int_0^{+\infty} x^2G(x)f_l''(x)dx < \infty$

A3. $\int_0^{+\infty} x^{-1}G^2(x)f_l(x)dx < \infty$.

Proposition 3.1: 1 Assuming that conditions **C1**, **C2** are satisfied, and the integrability condition $\mathbb{E}[\int_0^x K_{LogN(lny,4ln(1+h_m))}dy] < \infty, \forall m$, the following holds

$$\hat{G}_{LogN}(x) \xrightarrow{a.s.} G(x), \quad (7)$$

where $\xrightarrow{a.s.}$ indicates the almost sure convergence.

Proof 3.1: *2*First, we have

$$\begin{aligned}\mathbb{E}[\hat{G}_{LogN}(x)] &= \mathbb{E}\int_0^x K_{LogN(\ln y,4\ln(1+h_m))}(Y)dy \\ &= \int_0^{+\infty}\int_0^x K_{LogN(\ln y,4\ln(1+h_m))}(u)g(u)dydu \\ &= \int_0^x\int_0^{+\infty} K_{LogN(\ln y,4\ln(1+h_m))}(u)g(u)dudy = \int_0^x \mathbb{E}[g(Z_{y,h})]dy,\end{aligned} \quad (8)$$

where $Z_{y,h}: LogN(\mu,\sigma)$ and $[g(Z_{y,h})] = g(y) + h[2yg'(y) + 2y^2g''(y)] + o(h)$, see Jin and Kawczak [6]. By replacing in (8), then we have

$$\begin{aligned}\mathbb{E}[\hat{G}_{LogN}(x)] &= \int_0^x \{g(y) + h[2yg'(y) + 2y^2g''(y)] + o(h)\} \\ &= \int_0^x g(y)dy + 2h\int_0^x \{yg'(y) + y^2g''(y)\}dy + o(h) \\ &= G(x) + 2h\{G(x) + x^2g'(x) - xg'(x)\} + o(h) \\ &= G(x) + O(h)\end{aligned} \quad (9)$$

Second, note that $\int_0^x K_{LogN(\ln y,4\ln(1+h_m))}(Y_j)dy$ are i.i.d and $\mathbb{E}[K_{LogN(\ln y,4\ln(1+h_m))}(Y)] < \infty, \forall m$. Therefore, using the strong rule of large numbers, we get

$$\hat{G}_{LogN}(x) - \mathbb{E}[K_{LogN(\ln y,4\ln(1+h_m))}(Y_j)] \xrightarrow{a.s.} 0 \quad as \quad m \to +\infty. \quad (10)$$

Then, using (9), (10) and the classical decomposition as follows:

$$\hat{G}_{LogN}(x) - G(x) = \left[\mathbb{E}(\hat{G}_{LogN}(x)) - G(x)\right] + \left[\hat{G}_{LogN}(x) - \mathbb{E}(\hat{G}_{LogN}(x))\right]$$

we complete the proof of the proposition.

Theorem 3.2: Let $\hat{R}_s$ be the estimator of the reliability parameter R_s with LogN kernel. Under conditions **C1**, **C2**, A1, A2 and A3, the asymptotic expressions for the bias, variance and MSE of $\hat{R}_s$ are given as follows:

(i) Bias

$$Bias(\hat{R}_s) = h_n \sum_{l=1}^{s} B(G, f_l) \prod_{j\neq l} R_j + o(h_n), \tag{11}$$

where

$$B(G, f_l) = \int_0^\infty G(x)\{2xf_{l'}(x) + 2x^2 f_{l''}(x)\}dx$$

(ii) Variance

$$Var(\hat{R}_s) = \frac{1}{nh_n^{1/2}} \sum_{l=1}^{s} V(G, f_l) \prod_{j\neq l} R_j^2 + o\left(\frac{1}{nh^{1/2}}\right), \tag{12}$$

where

$$V(G, f_l) = \frac{1}{4\sqrt{\pi}} \int_0^\infty \frac{G^2(x) f_l(x)}{x}\, dx.$$

(iii) Means Squared Errors (MSE)

$$\begin{aligned} MSE\{\hat{R}_s\} &= Bias^2\{\hat{R}_s\} + Var\{\hat{R}_s\} \\ &= h_n^2\left(\sum_{l=1}^{s} B(G, f_l) \prod_{j\neq l} R_j\right)^2 \\ &\quad + \frac{1}{nh_n^{1/2}} \sum_{l=1}^{s} V(G, f_l) \prod_{j\neq l} R_j^2 + o\left(h_n^2 + \frac{1}{nh^{1/2}}\right) \end{aligned} \tag{13}$$

Proof 3.2: The proof has been revised to be consistent with assumptions A1-A3 and the corrected formulation of Theorem 3.2. From the proposition 3.1, we can write

$$\hat{R}_s = \prod_{l=1}^{s} \int_0^{+\infty} \hat{f}_{l,LogN}(x) \hat{G}_{LogN}(x) dx = \prod_{l=1}^{s} \int_0^{+\infty} \hat{f}_{l,LogN}(x) G(x) dx \quad a.s.$$

For the demonstration, we consider the following notation

$$Z_{il} = \int_0^\infty G(x)\, K_{LogN(x,h)}(X_{il})\, dx, \quad \overline{Z}_l = \frac{1}{n}\sum_{i=1}^{n} Z_{il}, \quad and \quad \hat{R}_s = \prod_{e=1}^{s} \overline{Z}_l$$

Since the observations X_{il} are i.i.d. with density f_l, we have

$$\mathbb{E}(\overline{Z}_{\mathbb{l}}) = \mathbb{E}(Z_{\mathbb{l}}),$$

$$\mathbb{E}(Z_{\mathbb{l}}) = \int_0^\infty G(x)\left(\int_0^\infty K_{LogN(x,h_n)}(u)\, f_l(u)\, du\right)dx.$$

Using the asymptotic expansion of the log-normal kernel estimator J(in and Kawczak [6])), as $h_n \to 0$,

$$\int_0^\infty K_{LogN(x,h_n)}(u)\, f_l(u)\, du = f_l(x) + h\{2xf_l'(x) + 2x^2 f_l''(x)\} + o(h).$$

Substituting this expansion into the previous expression yields

$$\mathbb{E}(Z_e) = \int_0^\infty G(x)\, f_l(x)\, dx + h_n \int_0^\infty G(x)\{2xf_{l'}(x) + 2x^2 f_l''(x)\}dx + o(h_n).$$

Therefore,

$$\mathbb{E}(Z_l) = R_l + \delta_l + o(h_n),$$

where

$$\delta_l = h_n \int_0^\infty G(x)\{2xf_{l'}(x) + 2x^2 f_l''(x)\}dx.$$

Hence, the bias of $\overline{Z}_l$ is

$$Bias(\overline{Z}_l) = \mathbb{E}(\overline{Z}_l) - R_l = \delta_l = O(h_n).$$

Since

$$\mathbb{E}(\hat{R}_s) = \prod_{l=1}^{s} \mathbb{E}(\overline{Z}_l) = \prod_{l=1}^{s} (R_l + \delta_l),$$

the bias of $\hat{R}_s$ is

$$Bias(\hat{R}_s) = \prod_{l=1}^{s} (R_l + \delta_l) - \prod_{l=1}^{s} R_l.$$

We now apply a first-order Taylor expansion of the product around $(R_1, \dots, R_s)$. Since $\delta_l = O(h_n)$, we obtain

$$\prod_{l=1}^{s} (R_l + \delta_l) = \prod_{l=1}^{s} R_l + \sum_{l=1}^{s} \delta_l \prod_{j\neq l} R_j + O(\textstyle\sum_{l=1}^{s} \delta_l^2).$$

Substituting the expression of δ_l gives

$$Bias(\hat{R}_s) = \sum_{l=1}^{s} h_n\left(\prod_{j\neq l} R_j\right) \int_0^\infty G(x)\{2xf_{l'}(x) + 2x^2 f_l''(x)\}dx + o(h_n).$$

This completes the proof of the asymptotic bias.

Asymptotic variance.

We start from the decomposition

$$\overline{Z}_l = \mathbb{E}(Z_l) + (\overline{Z}_l - \mathbb{E}(Z_l)), \qquad \mathbb{E}(Z_l) = R_l + \delta_l,$$

and

$$\hat{R}_s - \mathbb{E}(\hat{R}_s) = \prod_{l=1}^{s} \overline{Z}_l - \prod_{l=1}^{s} (R_l + \delta_l).$$

Using the linear expansion of the product around $(R_l + \delta_l)$, we obtain

$$\prod_{l=1}^{s} (R_l + \delta_l + \xi_l) \approx \prod_{l=1}^{s} (R_l + \delta_l) + \sum_{l=1}^{s} \xi_l \prod_{j\neq l} (R_j + \delta_j),$$

where $\xi_l = \overline{Z}_l - \mathbb{E}(Z_l)$.

Therefore,

$$\hat{R}_s - \mathbb{E}(\hat{R}_s) \approx \sum_{l=1}^{s} \xi_l \prod_{j\neq l} (R_j + \delta_j).$$

Taking the variance on both sides yields

$$Var(\hat{R}_s) \approx Var\left(\sum_{l=1}^{s} \xi_l \prod_{j\neq l} (R_j + \delta_j)\right).$$

Since the variables $\overline{Z}_l$ are independent, the ξ_l are also independent, hence

$$Var(\hat{R}_s) \approx \sum_{l=1}^{s} Var(\xi_l) \prod_{j\neq l} (R_j + \delta_j)^2.$$

Moreover,

$$Var(\xi_l) = Var(\overline{Z}_l) = \frac{1}{n} Var(Z_l).$$

Now,

$$Var(Z_l) = \int_0^{\infty} \left(\int_0^{\infty} G(x) K_{LogN(x,h_n)}(u)dx\right)^2 f_l(u)du - R_l^2 \approx \frac{1}{4h_n^{1/2}\sqrt{\pi}} \int_0^{\infty} \frac{G^2(x)f(x)}{x}\, dx + o\left(\frac{1}{h_n^{1/2}}\right),$$

which we denote by

$$V(G, f_l) = \frac{1}{4\sqrt{\pi}} \int_0^{\infty} \frac{G^2(x)f(x)}{x}\, dx.$$

Hence,

$$Var(\overline{Z}_l) \approx h_n^{-1/2} n^{-1} V(G, f_l).$$

Substituting into the previous expression, we obtain the asymptotic variance

$$Var(\hat{R}_s) \approx h_n^{-1/2} n^{-1} \sum_{l=1}^{s} V(G, f_l) \prod_{j \neq l} R_j^2 + +o\left(\frac{1}{nh_n^{1/2}}\right).$$

By substituting the bias and variance of $\hat{R}_s$ expressions found in formulae (11) and (12), respectively, We can get the required result shown in formula (13).

3.3 Choice of bandwidth

In kernel estimation, the bandwidth choice is crucial. It depends on Choosing a bandwidth that minimizes the asymptotic mean squared error.

The formula for the asymptotic mean square error (AMSE) is

$$AMSE(\hat{R}_s) = h_n^2 \left(\sum_{l=1}^{s} B(G, f_l) \prod_{j \neq l} R_j\right)^2 + \frac{1}{nh_n^{1/2}} \sum_{l=1}^{s} V(G, f_l) \prod_{j \neq l} R_j^2. \quad (14)$$

By minimizing the AMSE given in (14) with respect to h_n, we obtain the optimal bandwidth for the estimator $\hat{R}_s$:

$$h_{opt} = \left[\frac{\sum_{l=1}^{s} V(G, f_l) \prod_{j \neq l} R_j^2}{4\left(\sum_{l=1}^{s} B(G, f_l) \prod_{j \neq l} R_j\right)^2}\right]^{2/5} n^{-2/5}. \quad (15)$$

It should be noted that the bandwidth ([15]) cannot be directly utilized in practice as it depends on the CDF G and the unknown density f_l, f_l', f_l''.

3.3.1 Rule of thumb approach

For the LogN kernel estimator, we develop the rule of thumb (RT) approach. This method is based on formula (15) and Silverman's concept [22]. We suggest substituting a known exponential reference model with the parameters λ_1 and λ_2 for the unknown density f_l and and CDF G, respectively.

In practice, the parameters λ_1 and λ_2 can be replaced by their estimates $\hat{\lambda}_1$ and $\hat{\lambda}_2$, obtained from the moment method.

$$\hat{\lambda}_1 = \frac{1}{\bar{X}}, \hat{\lambda}_2 = \frac{1}{\bar{Y}}$$

with

$$\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i \text{ and } \bar{Y} = \frac{1}{m} \sum_{j=1}^{m} Y_j$$

By substituting G and f_l in [15] with their exponential reference counterparts, we obtain the rule of thumb bandwidth

$$h_{RT} = \left[\frac{\sum_{l=1}^{s} V\left(G_{\exp(\hat{\lambda}_2)}, f_{\exp(\hat{\lambda}_1),l}\right) \prod_{j\neq l} R_{j,\exp}^2}{4n\left(\sum_{l=1}^{s} B\left(G_{\exp(\hat{\lambda}_2)}, f_{\exp(\hat{\lambda}_1),l}\right) \prod_{j\neq l} R_{j,\exp}\right)^2}\right]^{2/5} n^{-2/5}. \tag{16}$$

3.3.2 Unbiased cross validation approach

The optimal bandwidth is obtained by minimizing the UCV criterion defined by

$$(h_n, h_m) = \arg\min_{(h_n,h_m)} \mathrm{UCV}(h_n, h_m).$$

The squared error is written as

$$\mathrm{SE}(\hat{R}_s) = (\hat{R}_s - R_s)^2 = \hat{R}_s^2 - 2\hat{R}_s R_s + R_s^2.$$

Since the last term does not depend on the bandwidth parameters, we consider

$$\mathrm{CV}(h_n, h_m) = \hat{R}_s(\hat{R}_s - 2R_s).$$

Recall that

$$R_s = \prod_{l=1}^{s} R_l = \prod_{l=1}^{s} \mathbb{E}[G(X_l)], \qquad \hat{R}_s = \prod_{l=1}^{s} \hat{R}_l.$$

Thus,

$$\mathrm{CV}(h_n, h_m) = \hat{R}_s^2 - 2\hat{R}_s \prod_{l=1}^{s} \mathbb{E}[G(X_l)].$$

In practice, the expectation is replaced by its empirical estimator

$$\widehat{\mathbb{E}}[G(X_l)] = \frac{1}{n}\sum_{i=1}^{n} \hat{G}(X_{il}) = \frac{1}{nm}\sum_{i=1}^{n} \sum_{\substack{j=1\\ j\neq i}}^{m} \int_0^{X_{il}} K_{LogN(h_m,y)}(Y_{jl})\, dy.$$

Consequently, the UCV criterion is given by

$$\mathrm{UCV}(h_n, h_m) = \hat{R}_s\left(\hat{R}_s - 2\prod_{l=1}^{s} \widehat{\mathbb{E}}[G(X_l)]\right).$$

After algebraic simplifications, the UCV criterion can be expressed as

$$\begin{aligned}
&\mathrm{UCV}(h_n, h_m)\\
&= \frac{1}{(nm)^s}\prod_{l=1}^{s} \sum_{i=1}^{n} \sum_{j=1}^{m} \int_0^{+\infty} \int_0^{x} K_{LogN(x,h_n)}(X_{il}) K_{LogN(y,h_m)}(Y_{jl})\, dy\, dx\\
&\times \Big[\frac{1}{(nm)^s}\sum_{i=1}^{n} \sum_{j=1}^{m} \int_0^{+\infty} \int_0^{x} K_{LogN(x,h_n)}(X_{il}) K_{LogN(y,h_m)}(Y_{jl})\, dy\, dx\\
&- \frac{2}{\{n(m-1)\}^s}\sum_{i=1}^{n} \sum_{\substack{j=1\\ j\neq i}}^{m} \int_0^{X_{il}} K_{LogN(y_j,h_m)}(Y_t)\, dy\Big].
\end{aligned}$$

4. Simulation study

This simulation study evaluates the performance of the series stress-strength system estimator with a LogN kernel and four components ($s = 4$) in the system. The comparison focuses on the choice of smoothing parameter between the UCV and RT methods developed in section 3.3. Six nonnegative distributions are used for the monte carlo simulation study: Exponential (Exp), Lognormal (LogN), Gamma, Weibull, Burr, and Pareto distributions. For each target density, we assume the same sample size n for all strength samples. We generate 100 replications of samples for the following size pairs (n, m)=(50,50), (100,100), (200,200), (500,500), (100,200), (100,500), (200,100), and (500,100).

We consider that X_1, X_2, X_3, X_4 and Y are independent random variables and do not necessarily follow the same probability distribution. We then consider the following cases:

- Case I1. X_1, X_2, X_3, X_4: $Exp(1)$ and Y: $Exp(3)$,
- Case I2. X_1, X_2, X_3, X_4: $log.N(3,4)$ and Y: $logN(2,1)$,
- Case I3. X_1, X_2, X_3, X_4: $Gamma(3,1)$ and Y: $Gamma(1,1)$,
- Case I4. X_1: $Exp(1)$ X_2: $log.N(3,4)$ X_3: $Gamma(3,1)$ X_4: $Weibull(1/4,1/5)$ and Y: $Pareto(3,1)$,
- Case I5. X_1, X_2: $weibull(5,7)$ X_3: $Pareto(1,3)$ X_4: $Burr(1,1/2,1)$ and Y: $Gamma(1,1)$,
- Case I6. X_1: $logN(3,4)$ X_2: $gamma(2,1)$ X_3: $weibull(5,7)$ X_4: $Burr(1,1/2,1)$ and Y: $Exp(3)$,

To evaluate the performance of the proposed estimator, we employ the squared error (SE), defined as follows:

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

Table 1

Simulation results (SE) for the cases I1, I2 and I3 with UCV and RT methods, for $s = 4$.

	I1		**I2**		**I3**	
	R_s**=0.32**		R_s**=0.13**		R_s**=0.59**	
$n = m$	**RT**	**UCV**	**RT**	**UCV**	**RT**	**UCV**
(50, 50)	0.17665	0.02222	0.0233454	0.01722	0.53114	0.03145
(100, 100)	0.09839	0.01744	0.00524	0.00281	0.19723	0.00634
(200, 200)	0.02449	0.00346	0.00175	0.00248	0.04858	0.00571
(500, 500)	0.01043	0.00223	0.00106	0.00158	0.03730	0.00196
(100, 200)	0.08870	0.012151	0.01005	0.00588	0.19332	0.01057
(100, 500)	0.08592	0.00560	0.00622	0.00350	0.11224	0.00550
(200, 100)	0.03198	0.00478	0.00308	0.00280	0.06734	0.00866
(500, 100)	0.01297	0.00310	0.00285	0.00211	0.04197	0.00410

Table 2

Simulation results (SE) for the cases I4, I5 and I6 with UCV and RT methods, for $s = 4$.

	I4		I5		I6	
	R=0.27		R=0.43		R=0.55	
$n = m$	RT	UCV	RT	UCV	RT	UCV
(50,50)	0.48213	0.00806	0.41753	0.06604	0.76104	0.04940
(100,100)	0.24823	0.00638	0.20481	0.02075	0.37880	0.03549
(200,200)	0.12859	0.00537	0.12926	0.00755	0.23802	0.01899
(500,500)	0.07604	0.00423	0.07769	0.00564	0.07913	0.01495
(100,200)	0.20238	0.03893	0.28125	0.01318	0.37132	0.06261
(100,500)	0.17403	0.00974	0.22497	0.00655	0.33645	0.03499
(200,100)	0.12441	0.01051	0.17694	0.01844	0.44764	0.04648
(500,100)	0.09034	0.00595	0.10418	0.00496	0.09966	0.02569

Tables 1 and 2 present the mean SE values of the LogN kernel estimator for the models considered, based on 100 replications. The RT and UCV methods are used to select the smoothing parameter, aiming to identify the method that yields the best results. From these tables, it is evident that:

- The average SE values generally decrease with increasing sample size (n, m). Similar trends are observed when fixing n and varying m, and vice versa.
- According to the simulation results, the UCV bandwidth selection approach outperforms the RT approach cross the models examined, except in case I4 for sample sizes $(n, m) = (200, 200)$ and $(500, 500)$.
- The SE values for strengths that are part of the same distribution family are shown in Table 1, while Table 2 shows the SE values for strengths from different distribution families. It is notable that the SE values in Table 1 are lower (indicating better performance) compared to those in Table 2.

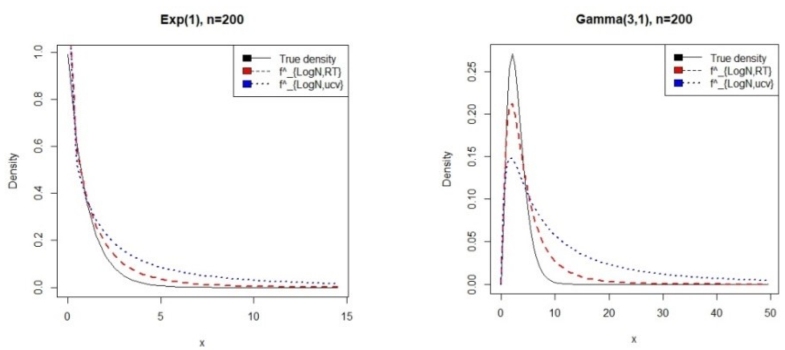

Figure 1
True strength density and Log-N kernel density estimator for gamma and exponential models with n = 200, using the bandwidths hRT and hUCV

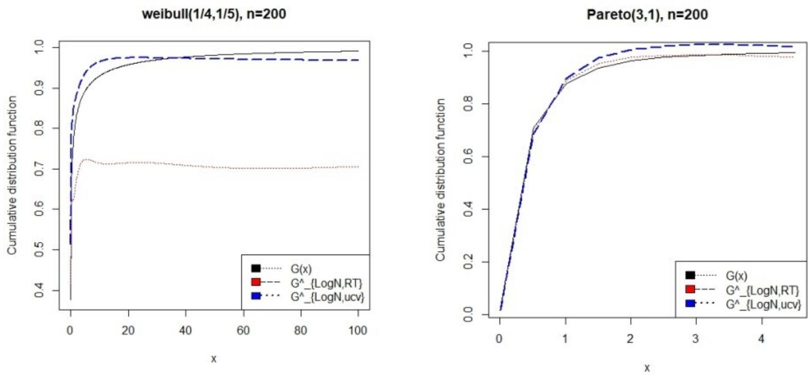

Figure 2
True stress CDF and Log-N kernel density estimator for Weibull and Pareto models with n = 200, using the bandwidths hRT and hUCV

The comparison is also illustrated in Figures 1 and 2. Figure 1 shows the true PDF and the Log-Normal kernel estimator of the strength density for the Gamma and Exponential models. Figure 2 shows the true CDF and the Log-N kernel estimator of the stress CDF for the Weibull and Pareto models, with a sample size of n=200. For the strength PDF estimator, the smoothing quality is very satisfactory when using the UCV and RT approachs with the Log-N kernel, and for the stress CDF estimator, the smoothing quality is very satisfactory when using the UCV method. This indicates that the proposed kernel provided good PDF and CDF estimations when combined with the RT and UCV approachs. The graphical analysis allows us to conclude that the Log-N kernel with the RT approach, provides an accurate estimation of the strength PDF, whereas its combination with the UCV approach provides a better estimation of the stress CDF.

5. Conclusion

In this paper, we proposed a nonparametric kernel estimator for series system stress strength reliability using a Log-Normal kernel, introduced by Jin and Kawczak [6]. Under some conditions, the estimator is asymptotically unbiased with variance and mean squared error converge to zero. A simulation study is conducted to evaluate the performance of the series system estimator using the unbiased cross validation and rule of thumb bandwidth selection methods. In general, the unbiased cross validation approach outperforms the rule of the thumb approach for selecting the bandwidth. Finally, examining different asymmetric kernels and various smoothing parameter selection methods would be intriguing. This study could serve as a foundation for extensive future research.

References

[1] M. Aslam, "Product acceptance determinations using new nonparametric sign sampling plan," *J. Stat. Manag. Syst.*, vol. 23, no. 8, pp. 1561-1570 (2020), doi.org/10.1080/09720510.2020.1764698.

[2] H. Ozen, E. Colak, C. Bal, F. Mutlu, and K. Ozdamar, "Is nonparametric bootstrap an appropriate technique for estimating variance of the sample median?," *Journal of Statistics and Management Systems*, vol. 24, no. 3, pp. 1–15 (2021), doi: 10.1080/09720510.2020.1866302.

[3] S. X. Chen, "Beta kernel estimators for density functions," *Comput. Stat. Data Anal.*, vol. 31, no. 2, pp. 131-145 (1999), doi.org/10.1016/S0167-9473(99)00010-9.

[4] S. X. Chen, "Probability density function estimation using gamma kernels," *Ann. Inst. Stat. Math.*, vol. 52, pp. 471-480 (2000), doi.org/10.1023/A:1004165218295.

[5] O. Scaillet, "Density estimation using inverse and reciprocal inverse Gaussian kernels," *Nonparametric Stat.*, vol. 16, no. 1-2, pp. 217-226 (2004), /doi.org/10.1080/10485250310001624819.

[6] X. Jin and J. Kawczak, "Birnbaum-Saunders and lognormal kernel estimators for modelling durations in high frequency financial data," *Annals of Economics and Finance*, vol. 4, no. 1, pp. 103–124 (2003).

[7] C. Marchant, K. Bertin, V. Leiva, and H. Saulo, "Generalized Birnbaum–Saunders kernel density estimators and an analysis of financial data," *Computational Statistics & Data Analysis*, vol. 63, pp. 1–15 (2013), doi: 10.1016/j.csda.2013.01.013.

[8] T. Bouezmarni, A. El Ghouch, and M. Mesfioui, "Gamma kernel estimators for density and hazard rate of right-censored data," *Journal of Probability and Statistics*, vol. 2011, Article ID 937574, pp. 1–16 (2011), doi: 10.1155/2011/937574.

[9] R.Salha, "Estimating the density and hazard rate functions using the reciprocal inverse Gaussian kernel," In 15th Applied Stochastic Models and Data Analysis International Conference., Mataro (Barcelona) Spain (2013, June).

[10] S. Chekkal, K. Lagha, and N. Zougab, "Generalized Birnbaum–Saunders kernel for hazard rate function estimation," *Communications in Statistics – Simulation and Computation*, vol. 52, no. 4, pp. 1546–1561 (2021), doi: 10.1080/03610918.2021.1887228.

[11] D. Kundu and R. D. Gupta, "Estimation of P(Y<X) for generalized exponential distribution," *Metrika.*, vol. 61, no. 3, pp. 291-308 (2005), doi.org/10.1007/s001840400345.

[12] D. Kundu and R. D. Gupta, "Estimation of P(Y<X) for Weibull distributions," *IEEE Trans. Reliab.,* vol. 55, no. 2, pp. 270-280 (2006).

[13] N. A. Mokhlis, "Reliability of a stress-strength model with Burr type III distributions," *Commun. Stat. Theory Methods.,* vol. 34, no. 7, pp. 1643-1657 (2005).

[14] M. Z. Raqab, M. T. Madi, and D. Kundu, "Estimation of P(Y<X) for the three-parameter generalized exponential distribution," *Commun. Stat. Theory Methods.,* vol. 37, no. 18, pp. 2854-2864 (2008), doi.org/10.1080/03610920802162664.

[15] S. Rezaei, R. Tahmasbi, and M. Mahmoodi, "Estimation of P(Y<X) for generalized Pareto distribution," *J. Stat. Plann. Inference.*, vol. 140, no. 2, pp. 480-494 (2010), doi.org/10.1016/j.jspi.2009.07.024.

[16] A. Asgharzadeh, R. Valiollahi, and M. Z. Raqab, "Estimation of the stress-strength reliability for the generalized logistic distribution," *Stat. Methodol.*, vol. 15, pp. 73-94 (2013), doi.org/10.1016/j.stamet.2013.05.002.

[17] G. S. Rao, R. R. L. Kantam, K. Rosaiah, and J. Pratapa Reddy, "Estimation of stress-strength reliability from inverse Rayleigh distribution," *J. Ind. Prod. Eng.,* vol. 30, no. 4, pp. 256-263 (2013), doi.org/10.1080/21681015.2013.828787.

[18] X. Jia, S. Nadarajah, and B. Guo, "Bayes estimation of P(Y<X) for the Weibull distribution with arbitrary parameters," *Appl. Math. Model.,* vol. 47, pp. 249-259 (2017), doi.org/10.1016/j.apm.2017.03.020.

[19] I. Muhammad, X. Wang, C. Li, M. Yan, and M. Chang, "Estimation of the reliability of a stress–strength system from Poisson half-logistic distribution," *Entropy,* vol. 22, no. 11, Art. no. 1307 (2020), doi: 10.3390/e22111307.

[20] A. M. Almarashi, A. Algarni, and M. Nassar, "On estimation procedures of stress–strength reliability for Weibull distribution with application," *PLOS ONE,* vol. 15, no. 8, Art. no. e0237997 (2020), doi: 10.1371/journal.pone.0237997.

[21] N. S. Y. Temraz, "Inference on the stress–strength reliability with exponentiated generalized Marshall–Olkin-G distribution," *PLOS ONE*, vol. 18, no. 8, Art. no. e0280183 (2023), doi: 10.1371/journal.pone.0280183.

[22] B. W. Silverman, Density Estimation for Statistics and Data Analysis. London, U.K.: Chapman and Hall (1986).

Received November, 2024