

Article

A Study of Seven Asymmetric Kernels for the Estimation of Cumulative Distribution Functions

Pierre Lafaye de Micheaux ^{1,2,3}  and Frédéric Ouimet ^{4,5*} ¹ School of Mathematics and Statistics, UNSW Sydney, Sydney, NSW 2052, Australia; lafaye@unsw.edu.au² Desbrest Institute of Epidemiology and Public Health, INSERM and University Montpellier, 34093 Montpellier, France³ AMIS, Université Paul Valéry Montpellier 3, 34199 Montpellier, France⁴ Department of Mathematics and Statistics, McGill University, Montreal, QC H3A 0B9, Canada⁵ Division of Physics, Mathematics and Astronomy, California Institute of Technology, Pasadena, CA 91125, USA

* Correspondence: frederic.ouimet2@mcgill.ca

Abstract: In this paper, we complement a study recently conducted in a paper of H.A. Mombeni, B. Masouri and M.R. Akhoond by introducing five new asymmetric kernel c.d.f. estimators on the half-line $[0, \infty)$, namely the Gamma, inverse Gamma, LogNormal, inverse Gaussian and reciprocal inverse Gaussian kernel c.d.f. estimators. For these five new estimators, we prove the asymptotic normality and we find asymptotic expressions for the following quantities: bias, variance, mean squared error and mean integrated squared error. A numerical study then compares the performance of the five new c.d.f. estimators against traditional methods and the Birnbaum–Saunders and Weibull kernel c.d.f. estimators from Mombeni, Masouri and Akhoond. By using the same experimental design, we show that the LogNormal and Birnbaum–Saunders kernel c.d.f. estimators perform the best overall, while the other asymmetric kernel estimators are sometimes better but always at least competitive against the boundary kernel method from C. Tenreiro.

Keywords: asymmetric kernels; asymptotic statistics; nonparametric statistics; Gamma kernel; inverse Gamma kernel; LogNormal kernel; inverse Gaussian kernel; reciprocal inverse Gaussian kernel; Birnbaum–Saunders kernel; Weibull kernel

MSC: Primary: 62G05; Secondary: 60F05, 62G20



Citation: Lafaye de Micheaux, P.; Ouimet, F. A Study of Seven Asymmetric Kernels for the Estimation of Cumulative Distribution Functions. *Mathematics* **2021**, *9*, 2605. <https://doi.org/10.3390/math9202605>

Academic Editor: Antonio Di Crescenzo

Received: 15 September 2021

Accepted: 12 October 2021

Published: 16 October 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

In the context of density estimation, asymmetric kernel estimators were introduced by Aitchison and Lauder [1] on the simplex and studied theoretically for the first time by Chen [2] on $[0, 1]$ (using a Beta kernel), and by Chen [3] on $[0, \infty)$ (using a Gamma kernel). These estimators are designed so that the bulk of the kernel function varies with each point x in the support of the target density. More specifically, the parameters of the kernel function can vary in a way that makes the mode, the median or the mean equal to x . This variable smoothing allows asymmetric kernel estimators to behave better than traditional kernel estimators (see, e.g., Rosenblatt [4], Parzen [5]) near the boundary of the support in terms of their bias. Since the variable smoothing is integrated directly in the parametrization of the kernel function, asymmetric kernel estimators are also usually simpler to implement than boundary kernel methods (see, e.g., Gasser and Müller [6], Rice [7], Gasser et al. [8], Müller [9], Zhang and Karunamuni [10,11]). For these two reasons, asymmetric kernel estimators are, by now, well-known solutions to the boundary bias problem from which traditional kernel estimators suffer. In the past twenty years, various asymmetric kernels have been considered in the literature on density estimation:

- Beta kernel, when the target density is supported on $[0, 1]$, see, e.g., Chen [2], Bouezmarni and Rolin [12], Renault and Scaillet [13], Fernandes and Monteiro [14], Hirukawa [15], Bouezmarni and Rombouts [16], Zhang and Karunamuni [17], Bertin and Klutchnikoff [18,19], Igarashi [20];

- Gamma, inverse Gamma, LogNormal, inverse Gaussian, reciprocal inverse Gaussian, Birnbaum–Saunders and Weibull kernels, when the target density is supported on $[0, \infty)$, see, e.g., Chen [3], Jin and Kawczak [21], Scaillet [22], Bouezmarni and Scaillet [23], Fernandes and Monteiro [14], Bouezmarni and Rombouts [16,24,25], Igarashi and Kakizawa [26,27], Charpentier and Flachaire [28], Igarashi [29], Zougab and Adjabi [30], Kakizawa and Igarashi [31], Kakizawa [32], Zougab et al. [33], Zhang [34], Kakizawa [35];
- Dirichlet kernel, when the target density is supported on the d -dimensional unit simplex, see [1] and the first theoretical study by Ouimet and Tolosana-Delgado [36].
- Continuous associated kernels, the aim of which is to unify the theory of asymmetric kernels with the one for traditional kernels in both the univariate and multivariate settings, see, e.g., Kokonendji and Libengué Dobélé-Kpoka [37], Kokonendji and Somé [38,39].

The interested reader is referred to Hirukawa [40] and Section 2 of Ouimet and Tolosana-Delgado [36] for a review of some of these papers and an extensive list of papers dealing with asymmetric kernels in other settings.

In contrast, there are almost no papers dealing with the estimation of cumulative distribution functions (c.d.f.s) in the literature on asymmetric kernels. In fact, to the best of our knowledge, [41] seems to be the first (and only) paper in this direction if we exclude the closely related theory of Bernstein estimators. (In the setting of Bernstein estimators, c.d.f. estimation on compact sets was tackled, for example, by Babu et al. [42], Leblanc [43], Leblanc [44], Leblanc [45], Dutta [46], Jmaei et al. [47], Erdoğan et al. [48] and Wang et al. [49] in the univariate setting, and by Babu and Chaubey [50], Belalia [51], Dib et al. [52] and Ouimet [53,54] in the multivariate setting. In [55], the authors introduced Bernstein estimators with Poisson weights (also called Szasz estimators) for the estimation of c.d.f.s that are supported on $[0, \infty)$, see also Ouimet [56].)

In the present paper, we complement the study reported in [41] by introducing five new asymmetric kernel c.d.f. estimators, namely the Gamma, inverse Gamma, LogNormal, inverse Gaussian and reciprocal inverse Gaussian kernel c.d.f. estimators. Our goal is to prove several asymptotic properties for these five new estimators (bias, variance, mean squared error, mean integrated squared error and asymptotic normality) and compare their numerical performance against traditional methods and against the Birnbaum–Saunders and Weibull kernel c.d.f. estimators from [41]. As we will see in the discussion of the results (Section 9), the LogNormal and Birnbaum–Saunders kernel c.d.f. estimators perform the best overall, while the other asymmetric kernel estimators are sometimes better but always at least competitive against the boundary kernel method from [57].

1. The Models

Let $X_1, X_2, \dots, X_n$ be a sequence of i.i.d. observations from an unknown cumulative distribution function F supported on the half-line $[0, \infty)$. We consider the following seven asymmetric kernel estimators (the first five are new):

$$\hat{F}_{n,b}^{\text{Gam}}(x) := \frac{1}{n} \sum_{i=1}^n \bar{K}_{\text{Gam}}(X_i | b^{-1}x + 1, b), \quad (1)$$

$$\hat{F}_{n,b}^{\text{IGam}}(x) := \frac{1}{n} \sum_{i=1}^n \bar{K}_{\text{IGam}}(X_i | b^{-1} + 1, x^{-1}b), \quad (2)$$

$$\hat{F}_{n,b}^{\text{LN}}(x) := \frac{1}{n} \sum_{i=1}^n \bar{K}_{\text{LN}}(X_i | \log x, \sqrt{b}), \quad (3)$$

$$\hat{F}_{n,b}^{\text{IGau}}(x) := \frac{1}{n} \sum_{i=1}^n \bar{K}_{\text{IGau}}(X_i | x, b^{-1}x), \quad (4)$$

$$\hat{F}_{n,b}^{\text{RIG}}(x) := \frac{1}{n} \sum_{i=1}^n \bar{K}_{\text{RIG}}(X_i | x^{-1}(1-b)^{-1}, x^{-1}b^{-1}), \quad (5)$$

$$\hat{F}_{n,b}^{\text{B-S}}(x) := \frac{1}{n} \sum_{i=1}^n \bar{K}_{\text{B-S}}(X_i | x, \sqrt{b}), \quad (6)$$

$$\hat{F}_{n,b}^{\text{W}}(x) := \frac{1}{n} \sum_{i=1}^n \bar{K}_{\text{W}}(X_i | x/\Gamma(1+b), b^{-1}), \quad (7)$$

where $b > 0$ is a *smoothing* (or *bandwidth*) parameter, and

$$\begin{aligned}\bar{K}_{\text{Gam}}(t | \alpha, \theta) &:= \frac{\Gamma(\alpha, t/\theta)}{\Gamma(\alpha)}, & \alpha, \theta > 0, \\ \bar{K}_{\text{IGam}}(t | \alpha, \theta) &:= 1 - \frac{\Gamma(\alpha, 1/(t\theta))}{\Gamma(\alpha)}, & \alpha, \theta > 0, \\ \bar{K}_{\text{LN}}(t | \mu, \sigma) &:= 1 - \Phi\left(\frac{\log t - \mu}{\sigma}\right), & \mu, \sigma > 0, \\ \bar{K}_{\text{IGau}}(t | \mu, \lambda) &:= 1 - \Phi\left(\sqrt{\frac{\lambda}{t}}\left(\frac{t}{\mu} - 1\right)\right) - e^{2\lambda/\mu}\Phi\left(-\sqrt{\frac{\lambda}{t}}\left(\frac{t}{\mu} + 1\right)\right), & \mu, \lambda > 0, \\ \bar{K}_{\text{RIG}}(t | \mu, \lambda) &:= \Phi\left(\sqrt{\lambda t}\left(\frac{1}{t\mu} - 1\right)\right) + e^{2\lambda/\mu}\Phi\left(-\sqrt{\lambda t}\left(\frac{1}{t\mu} + 1\right)\right), & \mu, \lambda > 0, \\ \bar{K}_{\text{B-S}}(t | \beta, \alpha) &:= 1 - \Phi\left(\frac{1}{\alpha}\left(\sqrt{\frac{t}{\beta}} - \sqrt{\frac{\beta}{t}}\right)\right), & \beta, \alpha > 0, \\ \bar{K}_{\text{W}}(t | \lambda, k) &:= \exp\left(-\left(\frac{t}{\lambda}\right)^k\right), & \lambda, k > 0,\end{aligned}$$

denote, respectively, the survival function of the

- Gamma(α, θ) distribution (with the shape/scale parametrization);
- InverseGamma(α, θ) distribution (with the shape/scale parametrization);
- LogNormal(μ, σ) distribution;
- InverseGaussian(μ, λ) distribution;
- ReciprocalInverseGaussian(μ, λ) distribution;
- Birnbaum–Saunders(β, α) distribution;
- Weibull(λ, k) distribution.

The function $\Gamma(\alpha, z) := \int_z^\infty t^{\alpha-1}e^{-t}dt$ denotes the upper incomplete gamma function (where $\Gamma(\alpha) := \Gamma(\alpha, 0)$), and Φ denotes the c.d.f. of the standard normal distribution. The parametrizations are chosen so that

- The mode of the kernel function in (1) is x ;
- The median of the kernel function in (3) and (6) is x ;
- The mean of the kernel function in (2), (4), (5) and (7) is x .

In this paper, we will compare the numerical performance of the above seven asymmetric kernel c.d.f. estimators against the following three traditional estimators (K here is the c.d.f. of a kernel function):

$$\hat{F}_{n,b}^{\text{OK}}(x) := \frac{1}{n} \sum_{i=1}^n K\left(\frac{x - X_i}{b}\right), \quad (8)$$

$$\hat{F}_{n,b}^{\text{BK}}(x) := \frac{1}{n} \sum_{i=1}^n \left\{ K\left(\frac{x - X_i}{b}\right) \mathbb{1}_{[b,\infty)}(x) + K\left(\frac{x - X_i}{x}\right) \mathbb{1}_{(0,b)}(x) \right\}, \quad (9)$$

$$\hat{F}_n^{\text{EDF}}(x) := \frac{1}{n} \sum_{i=1}^n \mathbb{1}_{[X_i, \infty)}(x), \quad (10)$$

which denote, respectively, the ordinary kernel (OK) c.d.f. estimator (from Tiago de Oliveira [58], Nadaraja [59] or Watson and Leadbetter [60]), the boundary kernel (BK) c.d.f. estimator (from Example 2.3 in [57]) and the empirical c.d.f. (EDF) estimator.

2. Outline, Assumptions and Notation

2.1. Outline

In Sections 3–7, the asymptotic normality, and the asymptotic expressions for the bias, variance, mean squared error (MSE) and mean integrated squared error (MISE), are stated for the Gam, IGam, LN, IGau and RIG kernel c.d.f. estimators, respectively. The

proofs can be found in Appendices A–E, respectively. Aside from the asymptotic normality (which can easily be deduced), these results were obtained for the Birnbaum–Saunders and Weibull kernel c.d.f. estimators in [41]. In Section 8, we compare the performance of all seven asymmetric kernel estimators above with the three traditional estimators OK, BK and EDF, defined in (8)–(10). A discussion of the results and our conclusion follow in Sections 9 and 10. Technical calculations for the proofs of the asymptotic results are gathered in Appendix F.

2.2. Assumptions

Throughout the paper, we make the following two basic assumptions:

1. The target c.d.f. F has two continuous and bounded derivatives;
2. The smoothing (or bandwidth) parameter $b = b_n > 0$ satisfies $b \rightarrow 0$ as $n \rightarrow \infty$.

2.3. Notation

Throughout the paper, the notation $u = \mathcal{O}(v)$ means that $\limsup |u/v| < C < \infty$ as $n \rightarrow \infty$. The positive constant C can depend on the target c.d.f. F , but no other variable unless explicitly written as a subscript. For example, if C depends on a given point $x \in (0, \infty)$, we would write $u = \mathcal{O}_x(v)$. Similarly, the notation $u = o(v)$ means that $\lim |u/v| = 0$ as $n \rightarrow \infty$. Subscripts indicate which parameters the convergence rate can depend on. The symbol $\mathcal{D}$ over an arrow ' $\longrightarrow$ ' will denote the convergence in law (or distribution).

3. Asymptotic Properties of the c.d.f. Estimator with Gam Kernel

In this section, we find the asymptotic properties of the Gamma (Gam) kernel estimator defined in (1).

Lemma 1 (Bias and variance). *For any given $x \in (0, \infty)$,*

$$\text{Bias}[\hat{F}_{n,b}^{\text{Gam}}(x)] := \mathbb{E}[\hat{F}_{n,b}^{\text{Gam}}(x)] - F(x) = b \cdot (f(x) + \frac{x}{2}f'(x)) + o_x(b), \quad (11)$$

$$\text{Var}(\hat{F}_{n,b}^{\text{Gam}}(x)) = n^{-1}F(x)(1 - F(x)) - n^{-1}b^{1/2} \cdot \frac{\sqrt{x}f(x)}{\sqrt{\pi}} + \mathcal{O}_x(n^{-1}b). \quad (12)$$

Corollary 1 (Mean squared error). *For any given $x \in (0, \infty)$,*

$$\begin{aligned} \text{MSE}(\hat{F}_{n,b}^{\text{Gam}}(x)) &= \text{Var}(\hat{F}_{n,b}^{\text{Gam}}(x)) + (\text{Bias}[\hat{F}_{n,b}^{\text{Gam}}(x)])^2 \\ &= n^{-1}F(x)(1 - F(x)) - n^{-1}b^{1/2} \cdot \frac{\sqrt{x}f(x)}{\sqrt{\pi}} \\ &\quad + b^2 \cdot (f(x) + \frac{x}{2}f'(x))^2 + \mathcal{O}_x(n^{-1}b) + o_x(b^2). \end{aligned} \quad (13)$$

In particular, if $f(x) \cdot (f(x) + xf'(x)) \neq 0$, the asymptotically optimal choice of b , with respect to MSE, is

$$b_{\text{opt}} = n^{-2/3} \left[\frac{4 \cdot (f(x) + \frac{x}{2}f'(x))^2}{\sqrt{x}f(x)/\sqrt{\pi}} \right]^{-2/3} \quad (14)$$

with

$$\begin{aligned} \text{MSE}(\hat{F}_{n,b_{\text{opt}}}^{\text{Gam}}(x)) &= n^{-1}F(x)(1 - F(x)) \\ &\quad - n^{-4/3} \frac{3}{4} \left[\frac{(\sqrt{x}f(x)/\sqrt{\pi})^4}{4 \cdot (f(x) + \frac{x}{2}f'(x))^2} \right]^{1/3} + o_x(n^{-4/3}). \end{aligned} \quad (15)$$

Proposition 1 (Mean integrated squared error). *Assuming that the target density $f = F'$ satisfies*

$$\int_0^\infty \sqrt{x}f(x)dx < \infty \quad \text{and} \quad \int_0^\infty (f(x) + \frac{x}{2}f'(x))^2dx < \infty, \quad (16)$$

then we have

$$\begin{aligned} \text{MISE}(\hat{F}_{n,b}^{\text{Gam}}) &= \int_0^\infty \text{Var}(\hat{F}_{n,b}^{\text{Gam}}(x)) dx + \int_0^\infty (\text{Bias}[\hat{F}_{n,b}^{\text{Gam}}(x)])^2 dx \\ &= n^{-1} \int_0^\infty F(x)(1-F(x)) dx - n^{-1} b^{1/2} \int_0^\infty \frac{\sqrt{x}f(x)}{\sqrt{\pi}} dx \\ &\quad + b^2 \int_0^\infty (f(x) + \frac{x}{2}f'(x))^2 dx + o(n^{-1}b^{1/2}) + o(b^2). \end{aligned} \quad (17)$$

In particular, if $f(x) \cdot (f(x) + x f'(x)) \neq 0$, the asymptotically optimal choice of b , with respect to MISE, is

$$b_{\text{opt}} = n^{-2/3} \left[\frac{4 \int_0^\infty (f(x) + \frac{x}{2}f'(x))^2 dx}{\int_0^\infty \sqrt{x}f(x)/\sqrt{\pi} dx} \right]^{-2/3} \quad (18)$$

with

$$\begin{aligned} \text{MISE}(\hat{F}_{n,b_{\text{opt}}}^{\text{Gam}}) &= n^{-1} \int_0^\infty F(x)(1-F(x)) dx \\ &\quad - n^{-4/3} \frac{3}{4} \left[\frac{(\int_0^\infty \sqrt{x}f(x)/\sqrt{\pi} dx)^4}{4 \int_0^\infty (f(x) + \frac{x}{2}f'(x))^2 dx} \right]^{1/3} + o(n^{-4/3}). \end{aligned} \quad (19)$$

Proposition 2 (Asymptotic normality). For any $x > 0$ such that $0 < F(x) < 1$, we have the following convergence in distribution:

$$n^{1/2}(\hat{F}_{n,b}^{\text{Gam}}(x) - \mathbb{E}[\hat{F}_{n,b}^{\text{Gam}}(x)]) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \sigma^2(x)), \quad \text{as } b \rightarrow 0, n \rightarrow \infty, \quad (20)$$

where $\sigma^2(x) := F(x)(1-F(x))$. In particular, Lemma 1 implies

$$n^{1/2}(\hat{F}_{n,b}^{\text{Gam}}(x) - F(x)) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \sigma^2(x)), \quad \text{if } n^{1/2}b \rightarrow 0, \quad (21)$$

$$n^{1/2}(\hat{F}_{n,b}^{\text{Gam}}(x) - F(x)) \xrightarrow{\mathcal{D}} \mathcal{N}(\lambda \cdot (f(x) + \frac{x}{2}f'(x)), \sigma^2(x)), \quad \text{if } n^{1/2}b \rightarrow \lambda, \quad (22)$$

for any constant $\lambda > 0$.

4. Asymptotic Properties of the c.d.f. Estimator with IGam Kernel

In this section, we find the asymptotic properties of the inverse Gamma (IGam) kernel estimator defined in (2).

Lemma 2 (Bias and variance). For any given $x \in (0, \infty)$,

$$\text{Bias}[\hat{F}_{n,b}^{\text{IGam}}(x)] := \mathbb{E}[\hat{F}_{n,b}^{\text{IGam}}(x)] - F(x) = b \cdot \frac{x^2}{2} f'(x) + o_x(b), \quad (23)$$

$$\text{Var}(\hat{F}_{n,b}^{\text{IGam}}(x)) = n^{-1} F(x)(1-F(x)) - n^{-1} b^{1/2} \cdot \frac{x f(x)}{\sqrt{\pi}} + \mathcal{O}_x(n^{-1}b). \quad (24)$$

Corollary 2 (Mean squared error). For any given $x \in (0, \infty)$,

$$\begin{aligned} \text{MSE}(\hat{F}_{n,b}^{\text{IGam}}(x)) &= \text{Var}(\hat{F}_{n,b}^{\text{IGam}}(x)) + (\text{Bias}[\hat{F}_{n,b}^{\text{IGam}}(x)])^2 \\ &= n^{-1} F(x)(1-F(x)) - n^{-1} b^{1/2} \cdot \frac{x f(x)}{\sqrt{\pi}} \\ &\quad + b^2 \cdot \frac{x^4}{4} (f'(x))^2 + \mathcal{O}_x(n^{-1}b) + o_x(b^2). \end{aligned} \quad (25)$$

In particular, if $f(x) \cdot f'(x) \neq 0$, the asymptotically optimal choice of b , with respect to MSE, is

$$b_{\text{opt}} = n^{-2/3} \left[\frac{4 \cdot \frac{x^4}{4} (f'(x))^2}{xf(x)/\sqrt{\pi}} \right]^{-2/3} \quad (26)$$

with

$$\begin{aligned} \text{MSE}(\hat{F}_{n,b_{\text{opt}}}^{\text{IGam}}(x)) &= n^{-1} F(x)(1 - F(x)) \\ &\quad - n^{-4/3} \frac{3}{4} \left[\frac{(xf(x)/\sqrt{\pi})^4}{4 \cdot \frac{x^4}{4} (f'(x))^2} \right]^{1/3} + o_x(n^{-4/3}). \end{aligned} \quad (27)$$

Proposition 3 (Mean integrated squared error). Assuming that the target density $f = F'$ satisfies

$$\int_0^\infty xf(x)dx < \infty \quad \text{and} \quad \int_0^\infty x^4(f'(x))^2dx < \infty, \quad (28)$$

then we have

$$\begin{aligned} \text{MISE}(\hat{F}_{n,b}^{\text{IGam}}) &= \int_0^\infty \text{Var}(\hat{F}_{n,b}^{\text{IGam}}(x))dx + \int_0^\infty (\mathbb{B}\text{ias}[\hat{F}_{n,b}^{\text{IGam}}(x)])^2dx \\ &= n^{-1} \int_0^\infty F(x)(1 - F(x))dx - n^{-1}b^{1/2} \int_0^\infty \frac{xf(x)}{\sqrt{\pi}}dx \\ &\quad + b^2 \int_0^\infty \frac{x^4}{4} (f'(x))^2dx + o(n^{-1}b^{1/2}) + o(b^2). \end{aligned} \quad (29)$$

In particular, if $\int_0^\infty x^4(f'(x))^2dx > 0$, the asymptotically optimal choice of b , with respect to MISE, is

$$b_{\text{opt}} = n^{-2/3} \left[\frac{4 \int_0^\infty \frac{x^4}{4} (f'(x))^2dx}{\int_0^\infty xf(x)/\sqrt{\pi}dx} \right]^{-2/3} \quad (30)$$

with

$$\begin{aligned} \text{MISE}(\hat{F}_{n,b_{\text{opt}}}^{\text{IGam}}) &= n^{-1} \int_0^\infty F(x)(1 - F(x))dx \\ &\quad - n^{-4/3} \frac{3}{4} \left[\frac{(\int_0^\infty xf(x)/\sqrt{\pi}dx)^4}{4 \int_0^\infty \frac{x^4}{4} (f'(x))^2dx} \right]^{1/3} + o(n^{-4/3}). \end{aligned} \quad (31)$$

Proposition 4 (Asymptotic normality). For any $x > 0$ such that $0 < F(x) < 1$, we have the following convergence in distribution:

$$n^{1/2}(\hat{F}_{n,b}^{\text{IGam}}(x) - \mathbb{E}[\hat{F}_{n,b}^{\text{IGam}}(x)]) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \sigma^2(x)), \quad \text{as } b \rightarrow 0, n \rightarrow \infty, \quad (32)$$

where $\sigma^2(x) := F(x)(1 - F(x))$. In particular, Lemma 2 implies

$$n^{1/2}(\hat{F}_{n,b}^{\text{IGam}}(x) - F(x)) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \sigma^2(x)), \quad \text{if } n^{1/2}b \rightarrow 0, \quad (33)$$

$$n^{1/2}(\hat{F}_{n,b}^{\text{IGam}}(x) - F(x)) \xrightarrow{\mathcal{D}} \mathcal{N}(\lambda \cdot \frac{x^2 f'(x)}{2}, \sigma^2(x)), \quad \text{if } n^{1/2}b \rightarrow \lambda, \quad (34)$$

for any constant $\lambda > 0$.

5. Asymptotic Properties of the c.d.f. Estimator with LN Kernel

In this section, we find the asymptotic properties of the LogNormal (LN) kernel estimator defined in (3).

Lemma 3 (Bias and variance). For any given $x \in (0, \infty)$,

$$\mathbb{B}\text{ias}[\hat{F}_{n,b}^{\text{LN}}(x)] := \mathbb{E}[\hat{F}_{n,b}^{\text{LN}}(x)] - F(x) = b \cdot \frac{x}{2}(f(x) + xf'(x)) + o_x(b), \quad (35)$$

$$\mathbb{V}\text{ar}(\hat{F}_{n,b}^{\text{LN}}(x)) = n^{-1}F(x)(1 - F(x)) - n^{-1}b^{1/2} \cdot \frac{xf(x)}{\sqrt{\pi}} + \mathcal{O}_x(n^{-1}b). \quad (36)$$

Corollary 3 (Mean squared error). For any given $x \in (0, \infty)$,

$$\begin{aligned} \text{MSE}(\hat{F}_{n,b}^{\text{LN}}(x)) &= \mathbb{V}\text{ar}(\hat{F}_{n,b}^{\text{LN}}(x)) + (\mathbb{B}\text{ias}[\hat{F}_{n,b}^{\text{LN}}(x)])^2 \\ &= n^{-1}F(x)(1 - F(x)) - n^{-1}b^{1/2} \cdot \frac{xf(x)}{\sqrt{\pi}} \\ &\quad + b^2 \cdot \frac{x^2}{4}(f(x) + xf'(x))^2 + \mathcal{O}_x(n^{-1}b) + o_x(b^2). \end{aligned} \quad (37)$$

In particular, if $f(x) \cdot (f(x) + xf'(x)) \neq 0$, the asymptotically optimal choice of b , with respect to MSE, is

$$b_{\text{opt}} = n^{-2/3} \left[\frac{4 \cdot \frac{x^2}{4}(f(x) + xf'(x))^2}{xf(x)/\sqrt{\pi}} \right]^{-2/3} \quad (38)$$

with

$$\begin{aligned} \text{MSE}(\hat{F}_{n,b_{\text{opt}}}^{\text{LN}}(x)) &= n^{-1}F(x)(1 - F(x)) \\ &\quad - n^{-4/3} \frac{3}{4} \left[\frac{(xf(x)/\sqrt{\pi})^4}{4 \cdot \frac{x^2}{4}(f(x) + xf'(x))^2} \right]^{1/3} + o_x(n^{-4/3}). \end{aligned} \quad (39)$$

Proposition 5 (Mean integrated squared error). Assuming that the target density $f = F'$ satisfies

$$\int_0^\infty xf(x)dx < \infty \quad \text{and} \quad \int_0^\infty x^2(f(x) + xf'(x))^2dx < \infty, \quad (40)$$

then we have

$$\begin{aligned} \text{MISE}(\hat{F}_{n,b}^{\text{LN}}) &= \int_0^\infty \mathbb{V}\text{ar}(\hat{F}_{n,b}^{\text{LN}}(x))dx + \int_0^\infty (\mathbb{B}\text{ias}[\hat{F}_{n,b}^{\text{LN}}(x)])^2dx \\ &= n^{-1} \int_0^\infty F(x)(1 - F(x))dx - n^{-1}b^{1/2} \int_0^\infty \frac{xf(x)}{\sqrt{\pi}}dx \\ &\quad + b^2 \int_0^\infty \frac{x^2}{4}(f(x) + xf'(x))^2dx + o(n^{-1}b^{1/2}) + o(b^2). \end{aligned} \quad (41)$$

In particular, if $f(x) \cdot (f(x) + xf'(x)) \neq 0$, the asymptotically optimal choice of b , with respect to MISE, is

$$b_{\text{opt}} = n^{-2/3} \left[\frac{4 \int_0^\infty \frac{x^2}{4}(f(x) + xf'(x))^2dx}{\int_0^\infty xf(x)/\sqrt{\pi}dx} \right]^{-2/3} \quad (42)$$

with

$$\begin{aligned} \text{MISE}(\hat{F}_{n,b_{\text{opt}}}^{\text{LN}}) &= n^{-1} \int_0^\infty F(x)(1 - F(x))dx \\ &\quad - n^{-4/3} \frac{3}{4} \left[\frac{(\int_0^\infty xf(x)/\sqrt{\pi}dx)^4}{4 \int_0^\infty \frac{x^2}{4}(f(x) + xf'(x))^2dx} \right]^{1/3} + o(n^{-4/3}). \end{aligned} \quad (43)$$

Proposition 6 (Asymptotic normality). For any $x > 0$ such that $0 < F(x) < 1$, we have the following convergence in distribution:

$$n^{1/2}(\hat{F}_{n,b}^{\text{LN}}(x) - \mathbb{E}[\hat{F}_{n,b}^{\text{LN}}(x)]) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \sigma^2(x)), \quad \text{as } b \rightarrow 0, n \rightarrow \infty, \quad (44)$$

where $\sigma^2(x) := F(x)(1 - F(x))$. In particular, Lemma 3 implies

$$n^{1/2}(\hat{F}_{n,b}^{\text{LN}}(x) - F(x)) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \sigma^2(x)), \quad \text{if } n^{1/2}b \rightarrow 0, \quad (45)$$

$$n^{1/2}(\hat{F}_{n,b}^{\text{LN}}(x) - F(x)) \xrightarrow{\mathcal{D}} \mathcal{N}\left(\lambda \cdot \frac{x}{2}(f(x) + xf'(x)), \sigma^2(x)\right), \quad \text{if } n^{1/2}b \rightarrow \lambda, \quad (46)$$

for any constant $\lambda > 0$.

6. Asymptotic Properties of the c.d.f. Estimator with IGau Kernel

In this section, we find the asymptotic properties of the inverse Gaussian (IGau) kernel estimator defined in (4).

Lemma 4 (Bias and variance). *For any given $x \in (0, \infty)$,*

$$\mathbb{B}\text{ias}[\hat{F}_{n,b}^{\text{IGau}}(x)] := \mathbb{E}[\hat{F}_{n,b}^{\text{IGau}}(x)] - F(x) = b \cdot \frac{x^2}{2} f'(x) + o_x(b), \quad (47)$$

$$\begin{aligned} \mathbb{V}\text{ar}(\hat{F}_{n,b}^{\text{IGau}}(x)) &= n^{-1}F(x)(1 - F(x)) - n^{-1}b^{1/2} \cdot \frac{f(x)}{2} \left[\lim_{b \rightarrow 0} b^{-1/2} \mathbb{E}[|T_1 - T_2|] \right] \\ &\quad + \mathcal{O}_x(n^{-1}b), \end{aligned} \quad (48)$$

where $T_1, T_2 \stackrel{\text{i.i.d.}}{\sim} \text{IGau}(x, b^{-1}x)$.

Corollary 4 (Mean squared error). *For any given $x \in (0, \infty)$,*

$$\begin{aligned} \text{MSE}(\hat{F}_{n,b}^{\text{IGau}}(x)) &= \mathbb{V}\text{ar}(\hat{F}_{n,b}^{\text{IGau}}(x)) + (\mathbb{B}\text{ias}[\hat{F}_{n,b}^{\text{IGau}}(x)])^2 \\ &= n^{-1}F(x)(1 - F(x)) - n^{-1}b^{1/2} \cdot \frac{f(x)}{2} \left[\lim_{b \rightarrow 0} b^{-1/2} \mathbb{E}[|T_1 - T_2|] \right] \\ &\quad + b^2 \cdot \frac{x^4}{4} (f'(x))^2 + \mathcal{O}_x(n^{-1}b) + o_x(b^2), \end{aligned} \quad (49)$$

where $T_1, T_2 \stackrel{\text{i.i.d.}}{\sim} \text{IGau}(x, b^{-1}x)$. The quantity $\lim_{b \rightarrow 0} b^{-1/2} \mathbb{E}[|T_1 - T_2|]$ needs to be approximated numerically. In particular, if $f(x) \left[\lim_{b \rightarrow 0} b^{-1/2} \mathbb{E}[|T_1 - T_2|] \right] \cdot f'(x) \neq 0$, the asymptotically optimal choice of b , with respect to MSE, is

$$b_{\text{opt}} = n^{-2/3} \left[\frac{4 \cdot \frac{x^4}{4} (f'(x))^2}{\frac{f(x)}{2} \lim_{b \rightarrow 0} b^{-1/2} \mathbb{E}[|T_1 - T_2|]} \right]^{-2/3} \quad (50)$$

with

$$\begin{aligned} \text{MSE}(\hat{F}_{n,b_{\text{opt}}}^{\text{IGau}}(x)) &= n^{-1}F(x)(1 - F(x)) \\ &\quad - n^{-4/3} \frac{3}{4} \left[\frac{\left(\frac{f(x)}{2} \lim_{b \rightarrow 0} b^{-1/2} \mathbb{E}[|T_1 - T_2|] \right)^4}{4 \cdot \frac{x^4}{4} (f'(x))^2} \right]^{1/3} + o_x(n^{-4/3}). \end{aligned} \quad (51)$$

Proposition 7 (Mean integrated squared error). *Assuming that the target density $f = F'$ satisfies*

$$\int_0^\infty f(x) \left[\lim_{b \rightarrow 0} b^{-1/2} \mathbb{E}[|T_1 - T_2|] \right] dx < \infty \quad \text{and} \quad \int_0^\infty x^4 (f'(x))^2 dx < \infty, \quad (52)$$

where $T_1, T_2 \stackrel{\text{i.i.d.}}{\sim} \text{IGau}(x, b^{-1}x)$, then we have

$$\begin{aligned} \text{MISE}(\hat{F}_{n,b}^{\text{IGau}}) &= \int_0^\infty \text{Var}(\hat{F}_{n,b}^{\text{IGau}}(x)) dx + \int_0^\infty (\text{Bias}[\hat{F}_{n,b}^{\text{IGau}}(x)])^2 dx \\ &= n^{-1} \int_0^\infty F(x)(1-F(x)) dx \\ &\quad - n^{-1} b^{1/2} \int_0^\infty \frac{f(x)}{2} \left[\lim_{b \rightarrow 0} b^{-1/2} \mathbb{E}[|T_1 - T_2|] \right] dx \\ &\quad + b^2 \int_0^\infty \frac{x^4}{4} (f'(x))^2 dx + o(n^{-1} b^{1/2}) + o(b^2). \end{aligned} \quad (53)$$

The quantity $\lim_{b \rightarrow 0} b^{-1/2} \mathbb{E}[|T_1 - T_2|]$ needs to be approximated numerically. In particular, if $\int_0^\infty x^4 (f'(x))^2 dx > 0$, the asymptotically optimal choice of b , with respect to MISE, is

$$b_{\text{opt}} = n^{-2/3} \left[\frac{4 \int_0^\infty \frac{x^4}{4} (f'(x))^2 dx}{\int_0^\infty \frac{f(x)}{2} \lim_{b \rightarrow 0} b^{-1/2} \mathbb{E}[|T_1 - T_2|] dx} \right]^{-2/3} \quad (54)$$

with

$$\begin{aligned} \text{MISE}(\hat{F}_{n,b_{\text{opt}}}^{\text{IGau}}) &= n^{-1} \int_0^\infty F(x)(1-F(x)) dx \\ &\quad - n^{-4/3} \frac{3}{4} \left[\frac{\left(\int_0^\infty \frac{f(x)}{2} \lim_{b \rightarrow 0} b^{-1/2} \mathbb{E}[|T_1 - T_2|] dx \right)^4}{4 \int_0^\infty \frac{x^4}{4} (f'(x))^2 dx} \right]^{1/3} \\ &\quad + o(n^{-4/3}). \end{aligned} \quad (55)$$

Proposition 8 (Asymptotic normality). For any $x > 0$ such that $0 < F(x) < 1$, we have the following convergence in distribution:

$$n^{1/2}(\hat{F}_{n,b}^{\text{IGau}}(x) - \mathbb{E}[\hat{F}_{n,b}^{\text{IGau}}(x)]) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \sigma^2(x)), \quad \text{as } b \rightarrow 0, n \rightarrow \infty, \quad (56)$$

where $\sigma^2(x) := F(x)(1-F(x))$. In particular, Lemma 4 implies

$$n^{1/2}(\hat{F}_{n,b}^{\text{IGau}}(x) - F(x)) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \sigma^2(x)), \quad \text{if } n^{1/2}b \rightarrow 0, \quad (57)$$

$$n^{1/2}(\hat{F}_{n,b}^{\text{IGau}}(x) - F(x)) \xrightarrow{\mathcal{D}} \mathcal{N}\left(\lambda \cdot \frac{x^2}{2} f'(x), \sigma^2(x)\right), \quad \text{if } n^{1/2}b \rightarrow \lambda, \quad (58)$$

for any constant $\lambda > 0$.

7. Asymptotic Properties of the c.d.f. Estimator with RIG Kernel

In this section, we find the asymptotic properties of the reciprocal inverse Gaussian (RIG) kernel estimator defined in (5).

Lemma 5 (Bias and variance). For any given $x \in (0, \infty)$,

$$\text{Bias}[\hat{F}_{n,b}^{\text{RIG}}(x)] := \mathbb{E}[\hat{F}_{n,b}^{\text{RIG}}(x)] - F(x) = b \cdot \frac{x^2}{2} f'(x) + o_x(b), \quad (59)$$

$$\begin{aligned} \text{Var}(\hat{F}_{n,b}^{\text{RIG}}(x)) &= n^{-1} F(x)(1-F(x)) - n^{-1} b^{1/2} \cdot \frac{f(x)}{2} \left[\lim_{b \rightarrow 0} b^{-1/2} \mathbb{E}[|T_1 - T_2|] \right] \\ &\quad + O_x(n^{-1}b), \end{aligned} \quad (60)$$

where $T_1, T_2 \stackrel{\text{i.i.d.}}{\sim} \text{RIG}(x, b^{-1}x)$.

Corollary 5 (Mean squared error). *For any given $x \in (0, \infty)$,*

$$\begin{aligned} \text{MSE}(\hat{F}_{n,b}^{\text{RIG}}(x)) &= \text{Var}(\hat{F}_{n,b}^{\text{RIG}}(x)) + (\text{Bias}[\hat{F}_{n,b}^{\text{RIG}}(x)])^2 \\ &= n^{-1}F(x)(1-F(x)) - n^{-1}b^{1/2} \cdot \frac{f(x)}{2} \left[\lim_{b \rightarrow 0} b^{-1/2} \mathbb{E}[|T_1 - T_2|] \right] \\ &\quad + b^2 \cdot \frac{1}{4}x^4(f'(x))^2 + \mathcal{O}_x(n^{-1}b) + \mathcal{O}_x(b^2), \end{aligned} \quad (61)$$

where $T_1, T_2 \stackrel{\text{i.i.d.}}{\sim} \text{RIG}(x^{-1}(1-b)^{-1}, x^{-1}b^{-1})$. The quantity $\lim_{b \rightarrow 0} b^{-1/2} \mathbb{E}[|T_1 - T_2|]$ needs to be approximated numerically. In particular, if $f(x) \left[\lim_{b \rightarrow 0} b^{-1/2} \mathbb{E}[|T_1 - T_2|] \right] \cdot f'(x) \neq 0$, the asymptotically optimal choice of b , with respect to MSE, is

$$b_{\text{opt}} = n^{-2/3} \left[\frac{4 \cdot \frac{1}{4}x^4(f'(x))^2}{\frac{f(x)}{2} \lim_{b \rightarrow 0} b^{-1/2} \mathbb{E}[|T_1 - T_2|]} \right]^{-2/3} \quad (62)$$

with

$$\begin{aligned} \text{MSE}(\hat{F}_{n,b_{\text{opt}}}^{\text{RIG}}(x)) &= n^{-1}F(x)(1-F(x)) \\ &\quad - n^{-4/3} \frac{3}{4} \left[\frac{\left(\frac{f(x)}{2} \lim_{b \rightarrow 0} b^{-1/2} \mathbb{E}[|T_1 - T_2|] \right)^4}{4 \cdot \frac{1}{4}x^4(f'(x))^2} \right]^{1/3} + \mathcal{O}_x(n^{-4/3}). \end{aligned} \quad (63)$$

Proposition 9 (Mean integrated squared error). *Assuming that the target density $f = F'$ satisfies*

$$\int_0^\infty f(x) \left[\lim_{b \rightarrow 0} b^{-1/2} \mathbb{E}[|T_1 - T_2|] \right] dx < \infty \quad \text{and} \quad \int_0^\infty x^4(f'(x))^2 dx < \infty, \quad (64)$$

where $T_1, T_2 \stackrel{\text{i.i.d.}}{\sim} \text{RIG}(x^{-1}(1-b)^{-1}, x^{-1}b^{-1})$, then we have

$$\begin{aligned} \text{MISE}(\hat{F}_{n,b}^{\text{RIG}}(x)) &= \int_0^\infty \text{Var}(\hat{F}_{n,b}^{\text{RIG}}(x)) dx + \int_0^\infty (\text{Bias}[\hat{F}_{n,b}^{\text{RIG}}(x)])^2 dx \\ &= n^{-1} \int_0^\infty F(x)(1-F(x)) dx \\ &\quad - n^{-1}b^{1/2} \int_0^\infty \frac{f(x)}{2} \left[\lim_{b \rightarrow 0} b^{-1/2} \mathbb{E}[|T_1 - T_2|] \right] dx \\ &\quad + b^2 \int_0^\infty \frac{x^4}{4} (f'(x))^2 dx + \mathcal{O}(n^{-1}b^{1/2}) + \mathcal{O}(b^2). \end{aligned} \quad (65)$$

The quantity $\lim_{b \rightarrow 0} b^{-1/2} \mathbb{E}[|T_1 - T_2|]$ needs to be approximated numerically. In particular, if $\int_0^\infty x^4(f'(x))^2 dx > 0$, the asymptotically optimal choice of b , with respect to MISE, is

$$b_{\text{opt}} = n^{-2/3} \left[\frac{4 \int_0^\infty \frac{1}{4}x^4(f'(x))^2 dx}{\int_0^\infty \frac{f(x)}{2} \lim_{b \rightarrow 0} b^{-1/2} \mathbb{E}[|T_1 - T_2|] dx} \right]^{-2/3} \quad (66)$$

with

$$\begin{aligned} \text{MISE}(\hat{F}_{n,b_{\text{opt}}}^{\text{RIG}}(x)) &= n^{-1} \int_0^\infty F(x)(1-F(x)) dx \\ &\quad - n^{-4/3} \frac{3}{4} \left[\frac{\left(\int_0^\infty \frac{f(x)}{2} \lim_{b \rightarrow 0} b^{-1/2} \mathbb{E}[|T_1 - T_2|] dx \right)^4}{4 \int_0^\infty \frac{1}{4}x^4(f'(x))^2 dx} \right]^{1/3} \\ &\quad + \mathcal{O}(n^{-4/3}). \end{aligned} \quad (67)$$

Proposition 10 (Asymptotic normality). *For any $x > 0$ such that $0 < F(x) < 1$, we have the following convergence in distribution:*

$$n^{1/2}(\hat{F}_{n,b}^{\text{RIG}}(x) - \mathbb{E}[\hat{F}_{n,b}^{\text{IGau}}(x)]) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \sigma^2(x)), \quad \text{as } b \rightarrow 0, n \rightarrow \infty, \quad (68)$$

where $\sigma^2(x) := F(x)(1 - F(x))$. In particular, Lemma 5 implies

$$n^{1/2}(\hat{F}_{n,b}^{\text{RIG}}(x) - F(x)) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \sigma^2(x)), \quad \text{if } n^{1/2}b \rightarrow 0, \quad (69)$$

$$n^{1/2}(\hat{F}_{n,b}^{\text{RIG}}(x) - F(x)) \xrightarrow{\mathcal{D}} \mathcal{N}\left(\lambda \cdot \frac{x^2}{2} f'(x), \sigma^2(x)\right), \quad \text{if } n^{1/2}b \rightarrow \lambda, \quad (70)$$

for any constant $\lambda > 0$.

8. Numerical Study

As in [41], we generated $M = 1000$ samples of size $n = 256$ and $n = 1000$ from eight target distributions:

1. Burr (1, 3, 1), with the following parametrization for the density function:

$$f_1(x | \lambda, c, k) := \frac{ck}{\lambda} \left(\frac{x}{\lambda}\right)^{c-1} \left[1 + \left(\frac{x}{\lambda}\right)^c\right]^{-k-1} \mathbb{1}_{(0,\infty)}(x), \quad \lambda, c, k > 0; \quad (71)$$

2. Gamma (0.6, 2), with the following parametrization for the density function:

$$f_2(x | \alpha, \theta) := \frac{x^{\alpha-1} \exp(-\frac{x}{\theta})}{\theta^\alpha \Gamma(\alpha)} \mathbb{1}_{(0,\infty)}(x), \quad \alpha, \theta > 0; \quad (72)$$

3. Gamma (4, 2), with the following parametrization for the density function:

$$f_3(x | \alpha, \theta) := \frac{x^{\alpha-1} \exp(-\frac{x}{\theta})}{\theta^\alpha \Gamma(\alpha)} \mathbb{1}_{(0,\infty)}(x), \quad \alpha, \theta > 0; \quad (73)$$

4. GeneralizedPareto (0.4, 1, 0), with the following parametrization for the density function:

$$f_4(x | \xi, \sigma, \mu) := \frac{1}{\sigma} \left[1 + \xi \left(\frac{x-\mu}{\sigma}\right)\right]^{-\frac{1}{\xi}-1} \mathbb{1}_{(\mu,\infty)}(x), \quad \xi, \sigma > 0, \mu \in \mathbb{R}; \quad (74)$$

5. HalfNormal (1), with the following parametrization for the density function:

$$f_5(x | \sigma) := \sqrt{\frac{2}{\pi\sigma^2}} \exp\left(-\frac{x^2}{2\sigma^2}\right) \mathbb{1}_{(0,\infty)}(x), \quad \sigma > 0; \quad (75)$$

6. LogNormal (0, 0.75), with the following parametrization for the density function:

$$f_6(x | \mu, \sigma) := \frac{1}{x\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(\log x - \mu)^2}{2\sigma^2}\right) \mathbb{1}_{(0,\infty)}(x), \quad \mu \in \mathbb{R}, \sigma > 0; \quad (76)$$

7. Weibull (1.5, 1.5), with the following parametrization for the density function:

$$f_7(x | \lambda, k) := \frac{k}{\lambda} \left(\frac{x}{\lambda}\right)^{k-1} \exp\left(-\left(\frac{x}{\lambda}\right)^k\right) \mathbb{1}_{(0,\infty)}(x), \quad \lambda, k > 0; \quad (77)$$

8. Weibull (3, 2), with the following parametrization for the density function:

$$f_8(x | \lambda, k) := \frac{k}{\lambda} \left(\frac{x}{\lambda}\right)^{k-1} \exp\left(-\left(\frac{x}{\lambda}\right)^k\right) \mathbb{1}_{(0,\infty)}(x), \quad \lambda, k > 0. \quad (78)$$

For each of the eight target distributions ($i = 1, 2, \dots, 8$), each of the ten estimators ($j = 1, 2, \dots, 10$), each sample size ($n = 256, 1000$), and each sample ($k = 1, 2, \dots, M$), we calculated the integrated squared errors

$$\text{ISE}_{i,j,n}^{(k)} := \int_0^\infty (\hat{F}_{j,n}^{(k)}(x) - F_i(x))^2 dx, \quad (79)$$

where

1. $\hat{F}_{1,n}^{(k)}$ denotes the estimator $\hat{F}_{n,b_{\text{opt}}}^{\text{Gam}}$ from (1) applied to the k -th sample;
2. $\hat{F}_{2,n}^{(k)}$ denotes the estimator $\hat{F}_{n,b_{\text{opt}}}^{\text{IGam}}$ from (2) applied to the k -th sample;
3. $\hat{F}_{3,n}^{(k)}$ denotes the estimator $\hat{F}_{n,b_{\text{opt}}}^{\text{LN}}$ from (3) applied to the k -th sample;
4. $\hat{F}_{4,n}^{(k)}$ denotes the estimator $\hat{F}_{n,b_{\text{opt}}}^{\text{IGau}}$ from (4) applied to the k -th sample;
5. $\hat{F}_{5,n}^{(k)}$ denotes the estimator $\hat{F}_{n,b_{\text{opt}}}^{\text{RIG}}$ from (5) applied to the k -th sample;
6. $\hat{F}_{6,n}^{(k)}$ denotes the estimator $\hat{F}_{n,b_{\text{opt}}}^{\text{B-S}}$ from (6) applied to the k -th sample;
7. $\hat{F}_{7,n}^{(k)}$ denotes the estimator $\hat{F}_{n,b_{\text{opt}}}^{\text{W}}$ from (7) applied to the k -th sample;

(for $\hat{F}_{1,n}^{(k)}$ to $\hat{F}_{7,n}^{(k)}$, b_{opt} is optimal with respect to the MISE (see (18), (30), (42), (54) and (66)) and approximated under the assumption that the target distribution is Gamma($\hat{\alpha}_n^{(k)}, \hat{\theta}_n^{(k)}$), where $\hat{\alpha}_n^{(k)}$ and $\hat{\theta}_n^{(k)}$ are the maximum likelihood estimates for the k -th sample) and

8. $\hat{F}_{8,n}^{(k)}(x) := \frac{1}{n} \sum_{\ell=1}^n \text{Epa}\left(\frac{x - X_\ell^{(k)}}{b_{\text{LNO}}}\right)$, where
 - $\text{Epa}(u) := \left(\frac{1}{2} + \frac{3u}{4} - \frac{u^3}{4}\right) \cdot \mathbb{1}_{(-1,1)}(u) + \mathbb{1}_{[1,\infty)}(u)$ denotes the c.d.f. of the Epanechnikov kernel;
 - b_{LNO} is selected by minimizing the Leave-None-Out criterion from page 197 in [61];
9. $\hat{F}_{9,n}^{(k)}(x) := \frac{1}{n} \sum_{\ell=1}^n \left\{ \text{Epa}\left(\frac{x - X_\ell^{(k)}}{b_{\text{CV}}}\right) \mathbb{1}_{[b_{\text{CV}}, \infty)}(x) + \text{Epa}\left(\frac{x - X_\ell^{(k)}}{x}\right) \mathbb{1}_{(0, b_{\text{CV}})}(x) \right\}$ is the boundary modified kernel estimator from Example 2.3 in [57], where
 - $\text{Epa}(u) := \left(\frac{1}{2} + \frac{3u}{4} - \frac{u^3}{4}\right) \cdot \mathbb{1}_{(-1,1)}(u) + \mathbb{1}_{[1,\infty)}(u)$ denotes the c.d.f. of the Epanechnikov kernel;
 - b_{CV} is selected by minimizing the Cross-Validation criterion from page 180 in [57];
10. $\hat{F}_{10,n}^{(k)}(x) := \frac{1}{n} \sum_{\ell=1}^n \mathbb{1}_{\{X_\ell^{(k)} \leq x\}}$ is the empirical c.d.f. applied to the k -th sample.

Everywhere in our R code, we approximated the integrals on $(0, \infty)$ using the `integral` function from the R package `pracma` (the base function `integrate` had serious precision issues). Table 1 below shows the mean and standard deviation of the ISE's, i.e.,

$$\frac{1}{M} \sum_{k=1}^M \text{ISE}_{i,j,n}^{(k)} \quad \text{and} \quad \sqrt{\frac{1}{M-1} \sum_{k=1}^M \left(\text{ISE}_{i,j,n}^{(k)} - \frac{1}{M} \sum_{k'=1}^M \text{ISE}_{i,j,n}^{(k')} \right)^2}, \quad (80)$$

for the eight target distributions ($i = 1, 2, \dots, 8$), the ten estimators ($j = 1, 2, \dots, 10$) and the two sample sizes ($n = 256, 1000$). All the values presented in the table have been multiplied by 10^4 . In Table 2, we computed, for each target distribution and each sample size, the difference between the ISE means and the lowest ISE mean for the corresponding target distribution and sample size (i.e., the ISE means minus the ISE mean of the best estimator on the corresponding line). The totals of those differences are also calculated for each sample size on the two "total" lines. Figure 1 gives a better idea of the target distribution of ISE's by displaying the boxplot of the ISE's for every target distribution and every estimator, when the sample size is $n = 1000$. Finally, Figures 2–9 (one figure for each

of the eight target distributions) show a collection of ten c.d.f. estimates from each of the ten estimators when the sample size is 256.

Here are the results, which we discuss briefly in Section 9:

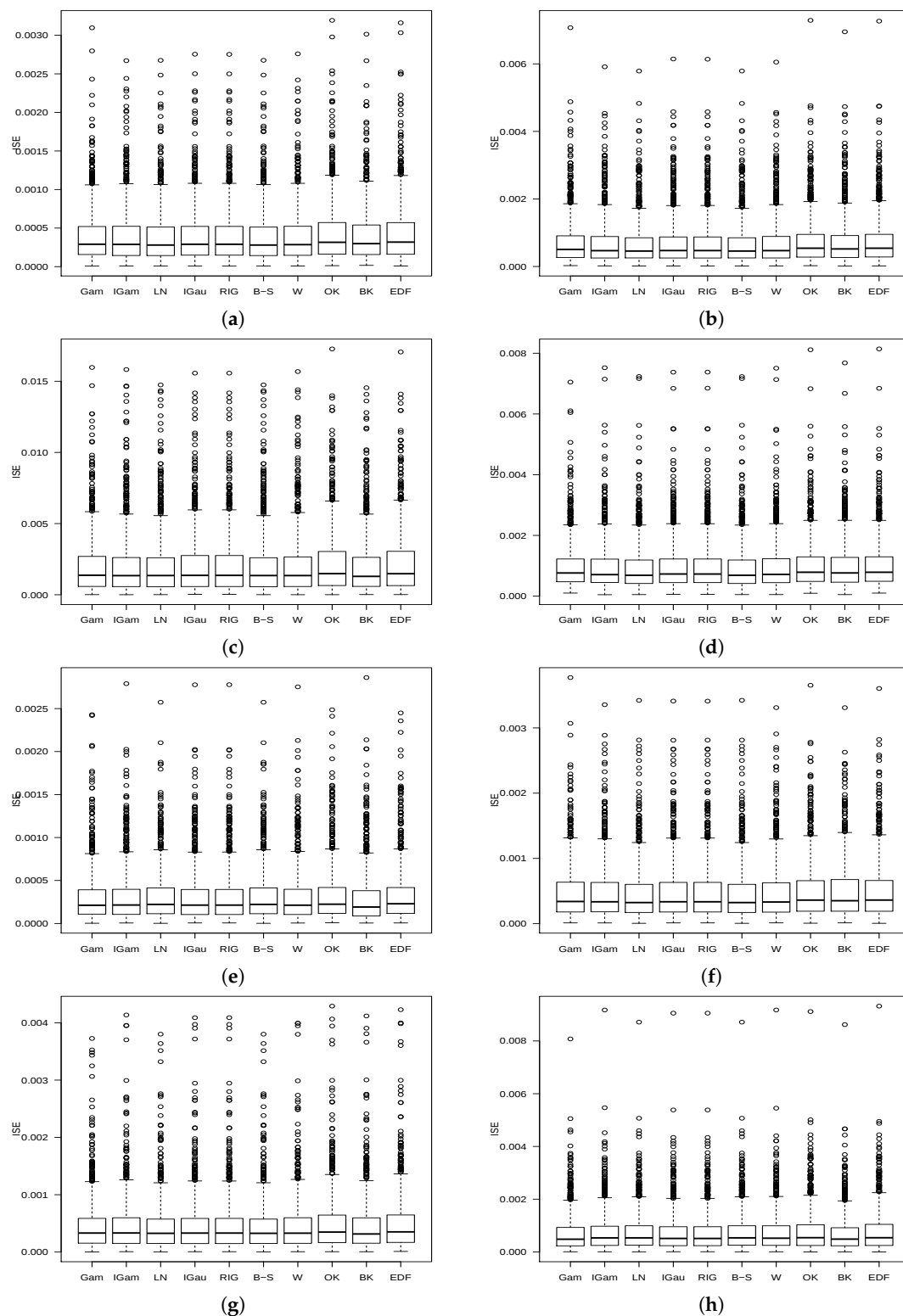


Figure 1. Boxplots of the $ISE_{i,j,n}^{(k)}$, $k = 1, 2, \dots, M$, for the eight target distributions and the ten estimators, when the sample size is $n = 1000$. (a) Burr(1, 3, 1). (b) Gamma(0.6, 2). (c) Gamma(4, 2). (d) GeneralizedPareto(0.4, 1, 0). (e) HalfNormal(1). (f) LogNormal(0, 0.75). (g) Weibull(1.5, 1.5). (h) Weibull(3, 2).

Table 1. The mean and standard deviation of the $ISE_{i,j,n}^{(k)}$, $k = 1, 2, \dots, M$, for the eight target distributions ($i = 1, 2, \dots, 8$), the ten estimators ($j = 1, 2, \dots, 10$) and the two sample sizes ($n = 256, 1000$). All the values presented in the table have been multiplied by 10^4 . The ordinary kernel estimator $\hat{F}_8$ is denoted by OK, the boundary kernel estimator $\hat{F}_9$ is denoted by BK, and the empirical c.d.f. $\hat{F}_{10}$ is denoted by EDF. For each line in the table, the lowest ISE means are highlighted in cyan.

$\times 10^{-4}$	i	Gam ($i = 1$)		IGam ($i = 2$)		LN ($i = 3$)		IGau ($i = 4$)		RIG ($i = 5$)		B-S ($i = 6$)		W ($i = 7$)		OK ($i = 8$)		BK ($i = 9$)		EDF ($i = 10$)	
n	j	mean	std.	mean	std.	mean	std.	mean	std.	mean	std.	mean	std.	mean	std.	mean	std.	mean	std.	mean	std.
256	1	1.39	1.27	1.37	1.34	1.31	1.26	1.37	1.32	1.37	1.32	1.31	1.26	1.37	1.32	1.54	1.43	1.47	1.34	1.54	1.44
	2	2.59	2.36	2.50	2.53	2.36	2.42	2.49	2.46	2.49	2.47	2.36	2.42	2.50	2.51	2.76	2.44	2.67	2.57	2.76	2.45
	3	6.70	6.28	6.77	6.58	6.62	6.28	6.69	6.45	6.69	6.45	6.62	6.28	6.74	6.54	7.44	7.01	6.70	6.39	7.44	7.00
	4	3.74	3.14	3.60	3.27	3.36	3.15	3.61	3.20	3.61	3.21	3.36	3.14	3.60	3.26	3.96	3.24	3.80	3.27	3.97	3.24
	5	1.14	1.10	1.18	1.13	1.18	1.07	1.17	1.13	1.17	1.13	1.18	1.07	1.17	1.12	1.26	1.19	1.10	1.14	1.26	1.19
	6	1.93	1.83	1.91	1.89	1.81	1.80	1.91	1.87	1.91	1.87	1.81	1.80	1.91	1.88	2.13	1.94	2.05	1.93	2.13	1.95
	7	1.75	1.82	1.77	1.99	1.68	1.83	1.76	1.96	1.76	1.96	1.68	1.83	1.76	1.95	1.95	1.93	1.73	2.04	1.95	1.92
	8	2.69	2.71	2.75	2.78	2.81	2.66	2.67	2.71	2.67	2.71	2.81	2.66	2.75	2.75	3.02	2.88	2.56	2.59	3.03	2.88
1000	1	0.40	0.36	0.39	0.36	0.38	0.35	0.39	0.36	0.39	0.36	0.38	0.35	0.39	0.36	0.43	0.39	0.41	0.36	0.43	0.39
	2	0.72	0.70	0.70	0.69	0.67	0.67	0.70	0.69	0.70	0.69	0.67	0.67	0.70	0.69	0.75	0.71	0.73	0.72	0.75	0.71
	3	2.01	2.09	2.05	2.22	2.02	2.16	2.04	2.15	2.04	2.15	2.02	2.16	2.05	2.20	2.23	2.29	1.99	2.09	2.23	2.30
	4	0.99	0.79	0.97	0.82	0.93	0.80	0.97	0.81	0.97	0.81	0.93	0.80	0.97	0.82	1.03	0.82	1.00	0.82	1.03	0.83
	5	0.31	0.31	0.31	0.31	0.31	0.30	0.31	0.31	0.31	0.31	0.31	0.30	0.31	0.31	0.33	0.32	0.30	0.32	0.33	0.32
	6	0.47	0.43	0.47	0.43	0.46	0.42	0.47	0.43	0.47	0.43	0.46	0.42	0.47	0.43	0.50	0.45	0.49	0.43	0.50	0.45
	7	0.46	0.46	0.46	0.48	0.44	0.45	0.46	0.48	0.46	0.48	0.44	0.45	0.46	0.48	0.49	0.50	0.46	0.48	0.49	0.50
	8	0.72	0.74	0.74	0.75	0.75	0.74	0.73	0.75	0.73	0.75	0.75	0.74	0.74	0.75	0.78	0.80	0.70	0.72	0.78	0.81

Table 2. For each of the eight target distributions ($i = 1, 2, \dots, 8$) and each of the two sample sizes ($n = 256, 1000$), a cell represents the mean of the $ISE_{i,j,n}^{(k)}$, $k = 1, 2, \dots, M$, minus the lowest ISE mean for that line (i.e., minus the ISE mean of the best estimator for that specific target distribution and sample size). For each estimator ($j = 1, 2, \dots, 10$) and each sample size, the total of those differences to the best ISE means is calculated on the line called “total”. For each sample size, the lowest totals are highlighted in cyan.

$\times 10^{-4}$	i	Gam ($i = 1$)	IGam ($i = 2$)	LN ($i = 3$)	IGau ($i = 4$)	RIG ($i = 5$)	B-S ($i = 6$)	W ($i = 7$)	OK ($i = 8$)	BK ($i = 9$)	EDF ($i = 10$)
n	j	diff. with Lowest Mean	diff. with Lowest Mean	diff. with Lowest Mean	diff. with Lowest Mean	diff. with Lowest Mean	diff. with Lowest Mean	diff. with Lowest Mean	diff. with Lowest Mean	diff. with Lowest Mean	diff. with Lowest Mean
256	1	0.08	0.06	0.00	0.06	0.06	0.00	0.06	0.23	0.16	0.23
	2	0.23	0.14	0.00	0.14	0.13	0.00	0.14	0.40	0.32	0.40
	3	0.08	0.15	0.01	0.08	0.07	0.00	0.12	0.82	0.09	0.82
	4	0.38	0.24	0.00	0.25	0.24	0.00	0.23	0.60	0.43	0.60
	5	0.05	0.08	0.09	0.07	0.07	0.08	0.08	0.16	0.00	0.17
	6	0.12	0.10	0.00	0.10	0.10	0.00	0.10	0.32	0.24	0.32
	7	0.07	0.10	0.00	0.09	0.09	0.00	0.08	0.27	0.05	0.27
	8	0.13	0.19	0.25	0.11	0.11	0.25	0.19	0.46	0.00	0.46
total		1.14	1.07	0.35	0.89	0.88	0.34	0.99	3.26	1.29	3.28
1000	1	0.02	0.01	0.00	0.01	0.01	0.00	0.01	0.05	0.03	0.05
	2	0.04	0.02	0.00	0.02	0.02	0.00	0.02	0.07	0.06	0.08
	3	0.02	0.06	0.03	0.05	0.05	0.03	0.06	0.24	0.00	0.24
	4	0.06	0.04	0.00	0.04	0.04	0.00	0.04	0.10	0.07	0.10
	5	0.01	0.02	0.02	0.01	0.01	0.02	0.02	0.03	0.00	0.03
	6	0.02	0.02	0.00	0.02	0.02	0.00	0.02	0.04	0.04	0.04
	7	0.01	0.02	0.00	0.02	0.02	0.00	0.02	0.05	0.02	0.05
	8	0.02	0.04	0.05	0.03	0.03	0.05	0.04	0.08	0.00	0.08
total		0.20	0.23	0.10	0.20	0.20	0.10	0.24	0.66	0.22	0.68

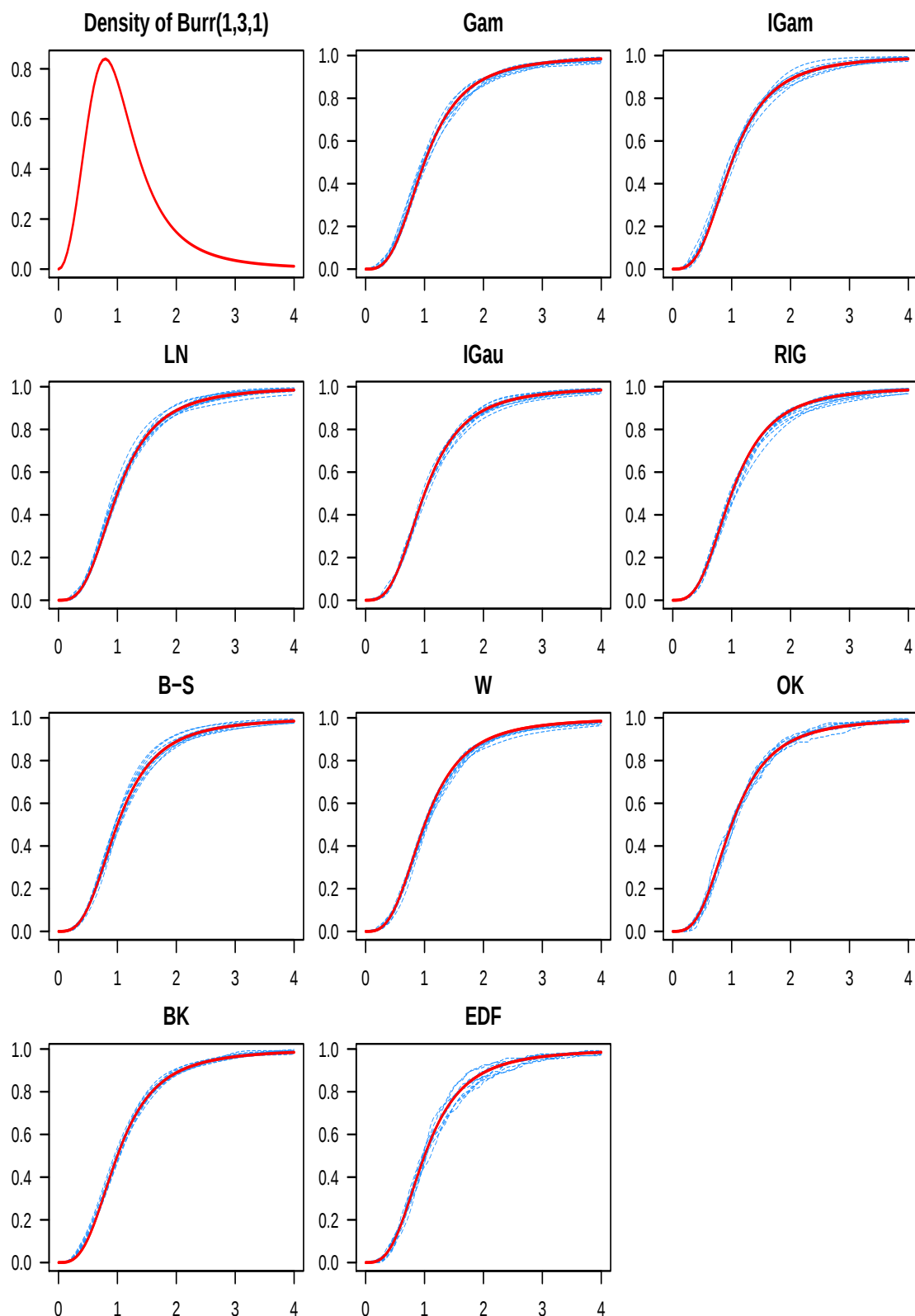


Figure 2. The Burr (1,3,1) density function appears on the top-left, and the target c.d.f. is depicted in red everywhere else. Each plot has ten estimates in blue for the Burr (1,3,1) c.d.f. using one of the ten estimators (the name of the corresponding kernel is indicated above each graph) and $n = 256$.

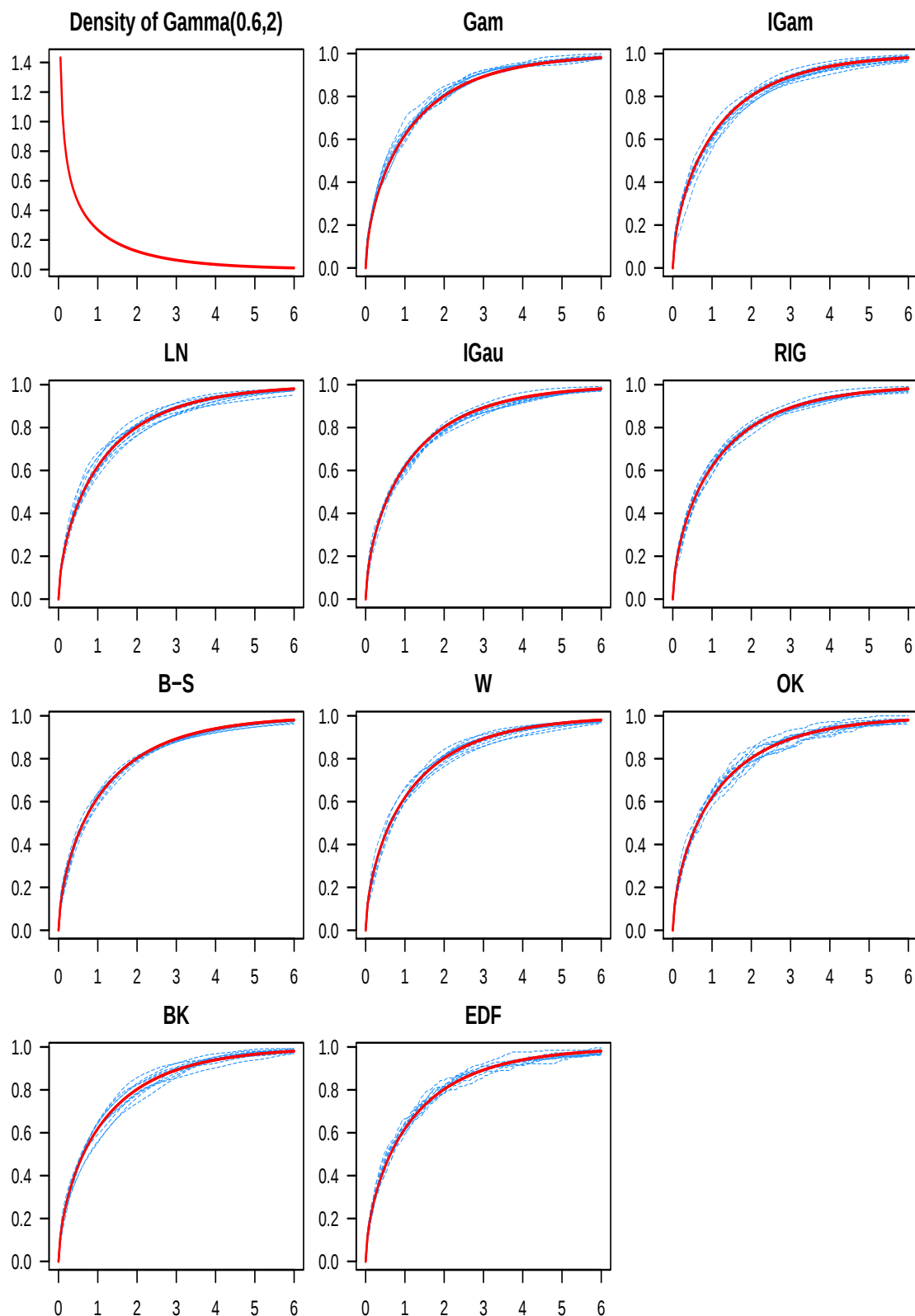


Figure 3. The Gamma (0.6, 2) density function appears on the top-left, and the target c.d.f. is depicted in red everywhere else. Each plot has ten estimates in blue for the Gamma (0.6, 2) c.d.f. using one of the ten estimators (the name of the corresponding kernel is indicated above each graph) and $n = 256$.

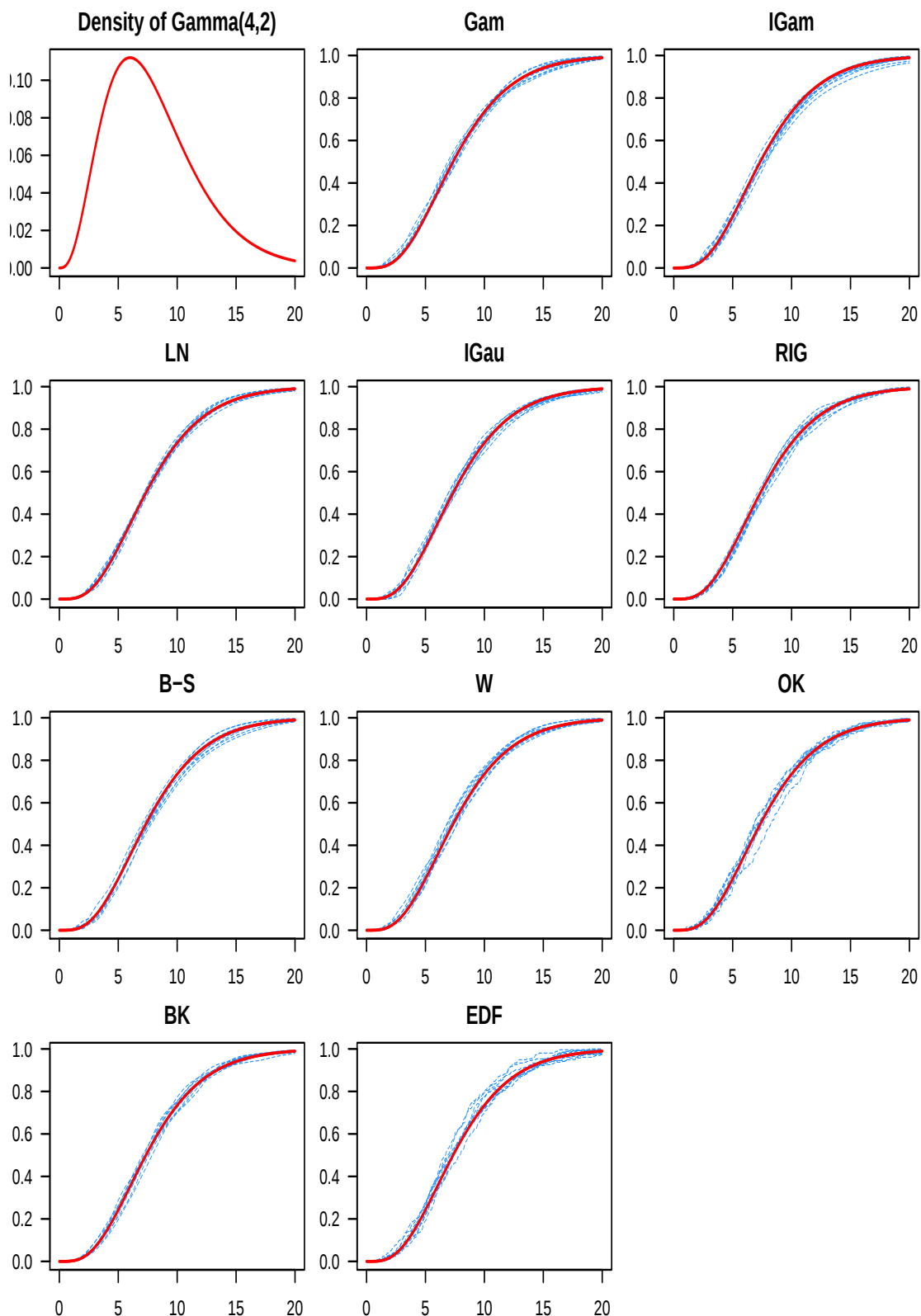


Figure 4. The Gamma (4,2) density function appears on the top-left, and the target c.d.f. is depicted in red everywhere else. Each plot has ten estimates in blue for the Gamma (4,2) c.d.f. using one of the ten estimators (the name of the corresponding kernel is indicated above each graph) and $n = 256$.

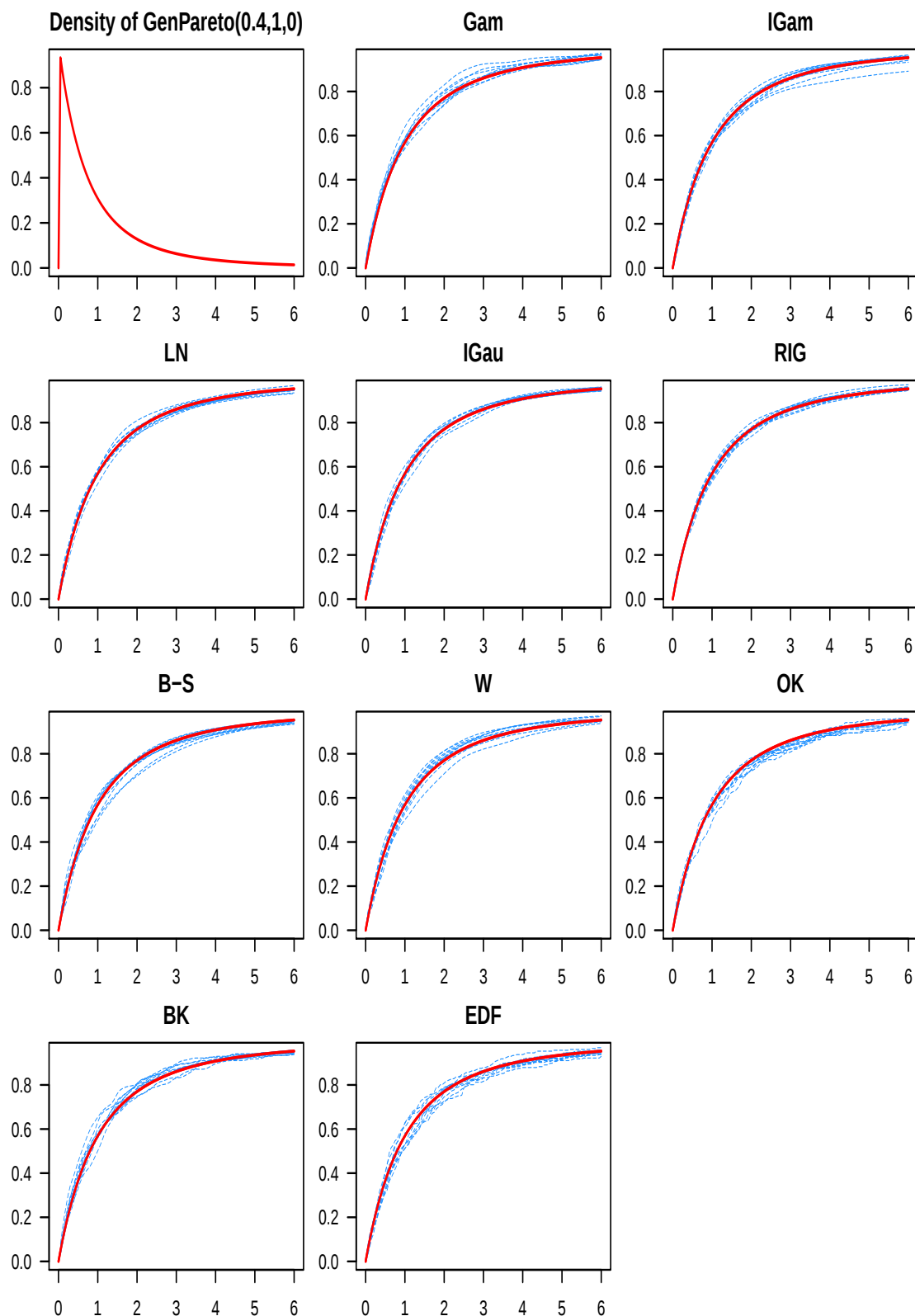


Figure 5. The GeneralizedPareto (0.4, 1, 0) density function appears on the top-left, and the target c.d.f. is depicted in red everywhere else. Each plot has ten estimates in blue for the GeneralizedPareto (0.4, 1, 0) c.d.f. using one of the ten estimators (the name of the corresponding kernel is indicated above each graph) and $n = 256$.

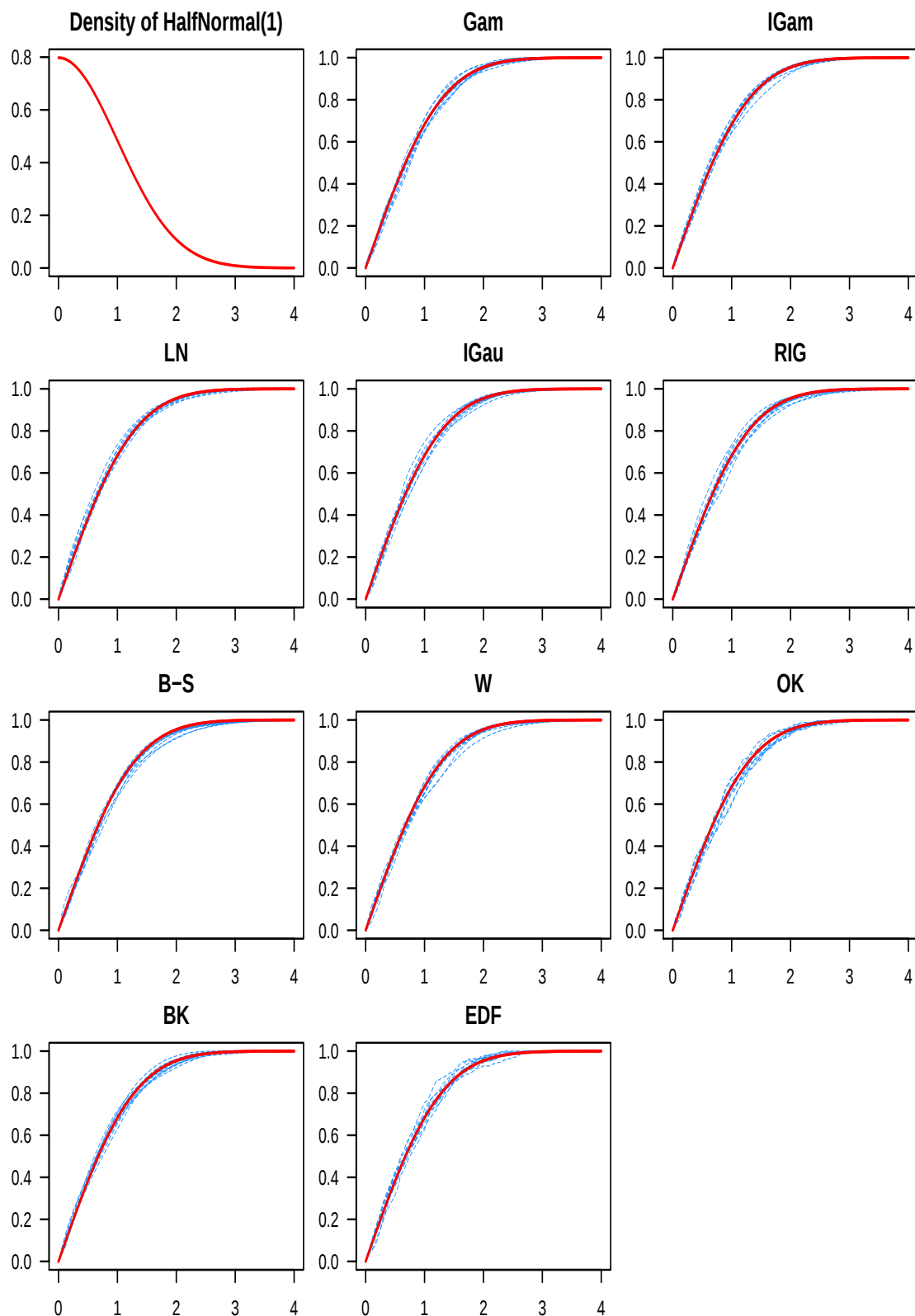


Figure 6. The HalfNormal (1) density function appears on the top-left, and the target c.d.f. is depicted in red everywhere else. Each plot has ten estimates in blue for the HalfNormal (1) c.d.f. using one of the ten estimators (the name of the corresponding kernel is indicated above each graph) and $n = 256$.

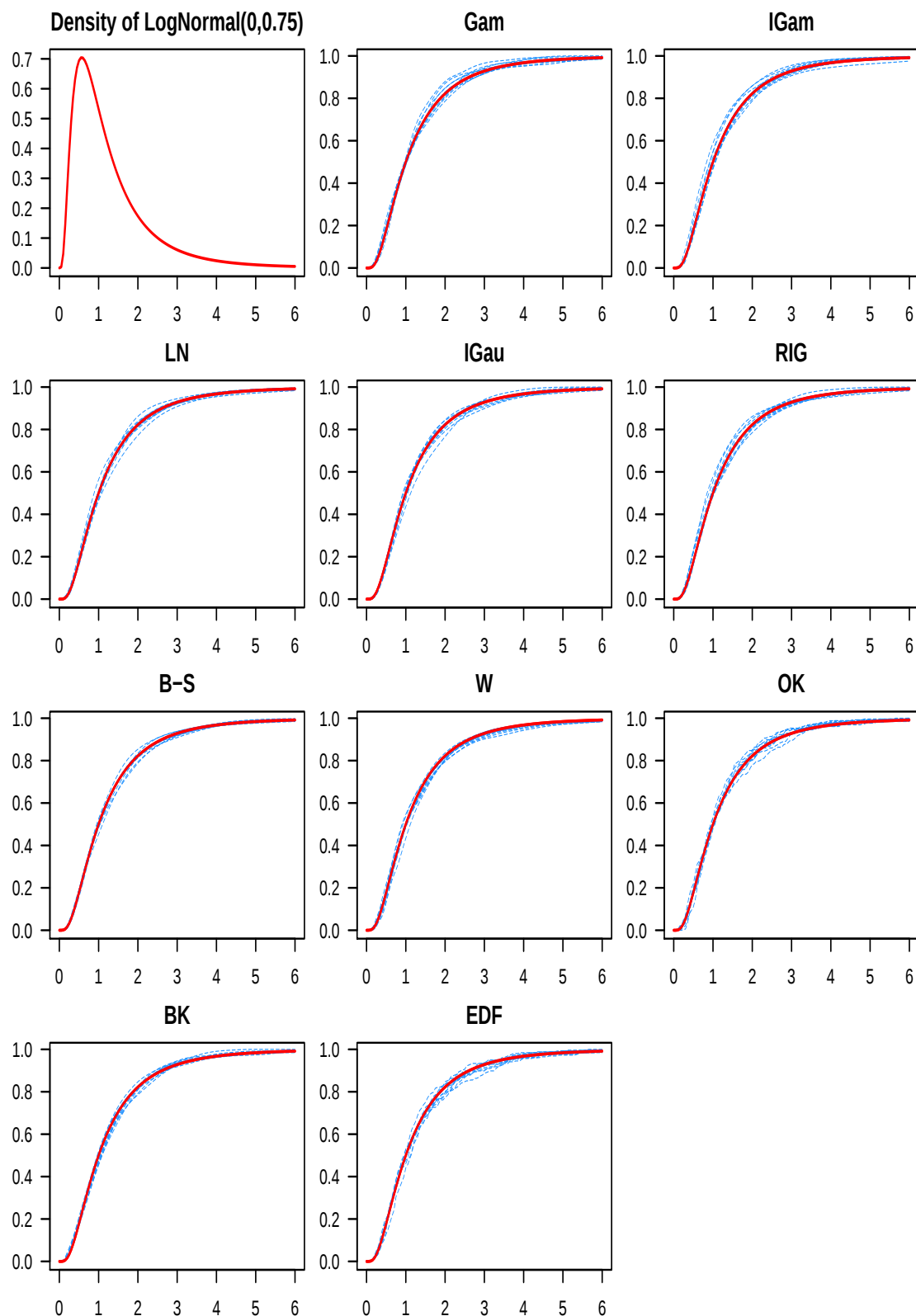


Figure 7. The LogNormal (0, 0.75) density function appears on the top-left, and the target c.d.f. is depicted in red everywhere else. Each plot has ten estimates in blue for the LogNormal (0, 0.75) c.d.f. using one of the ten estimators (the name of the corresponding kernel is indicated above each graph) and $n = 256$.

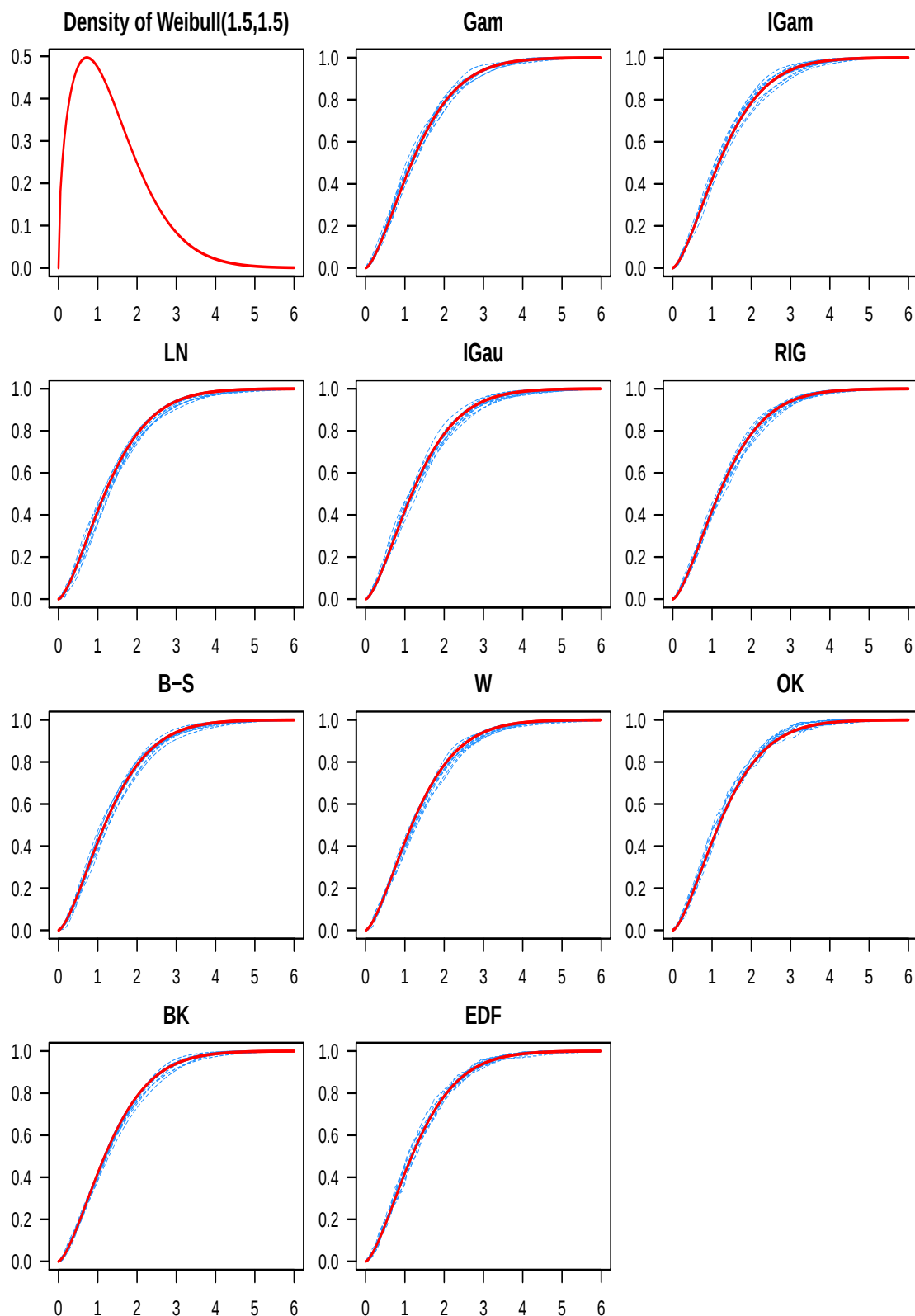


Figure 8. The Weibull (1.5, 1.5) density function appears on the top-left, and the target c.d.f. is depicted in red everywhere else. Each plot has ten estimates in blue for the Weibull (1.5, 1.5) c.d.f. using one of the ten estimators (the name of the corresponding kernel is indicated above each graph) and $n = 256$.

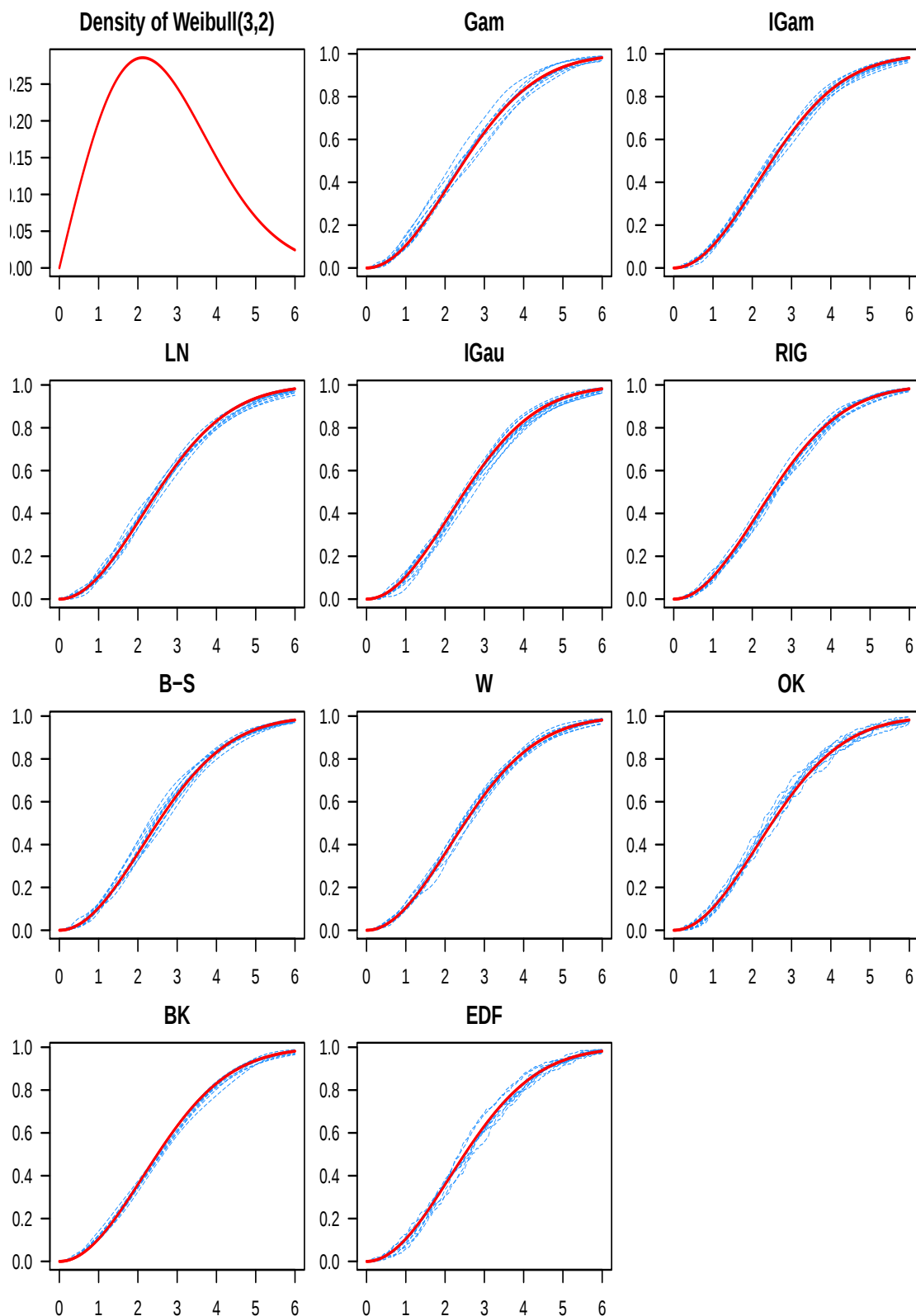


Figure 9. The Weibull (3,2) density function appears on the top-left, and the target c.d.f. is depicted in red everywhere else. Each plot has ten estimates in blue for the Weibull (3,2) c.d.f. using one of the ten estimators (the name of the corresponding kernel is indicated above each graph) and $n = 256$.

9. Discussion of the Simulation Results

In Table 1, the mean and standard deviation of the $ISE_{i,j,n}^{(k)}$, $k = 1, 2, \dots, M$, are displayed for the eight target distributions ($i = 1, 2, \dots, 8$), the ten estimators ($j = 1, 2, \dots, 10$) and the two sample sizes ($n = 256, 1000$). All the values presented in the table have been multiplied by 10^4 . For each line in the table (i.e., for each target distribution and each sample size), the lowest ISE means are highlighted in cyan. We see that the LogNormal (LN) and Birnbaum–Saunders (B–S) kernel c.d.f. estimators performed the best (had the lowest ISE means) for the majority of the target distributions considered (for $j = 1, 2, 3, 4, 6, 7$ when $n = 256$, and for $j = 1, 2, 4, 6, 7$ when $n = 1000$). They also always did so in pair, with the same ISE mean up to the second decimal. For the remaining cases, the boundary kernel c.d.f. estimator (BK) from Tenreiro [57] had the lowest ISE means. As expected, the ordinary kernel c.d.f. estimator and the empirical c.d.f. performed the worst. The standard deviations are fairly stable across all estimators for any given target distribution and sample size (this can also be seen in Figure 1), so our analysis focuses on the ISE means. In [41], the authors reported that the empirical c.d.f. performed better than the BK estimator, but this has to be a programming error (especially since the bandwidth was optimized with a plug-in method). Overall, our means and standard deviations in Table 1 seem to be lower than the ones reported in [41] at least in part because we used a more precise option (the `pracma::integral` function in R) to approximate the integrals involved in the bandwidth selection procedures and the computation of the ISE's. In all cases, the asymmetric kernel estimators were at least competitive with the BK estimator in Table 1. To give an idea of the shape of the eight target distributions and the corresponding estimates for each of the ten estimators, we plotted the eight target c.d.f.s and ten estimates for each estimator (one figure for each of the eight target distributions, ten graphs per figure for the ten estimators, and ten estimates per graph) in Figures 2–9 when the sample size is $n = 256$.

In Table 2, for each of the eight target distributions ($i = 1, 2, \dots, 8$) and each of the two sample sizes ($n = 256, 1000$), a cell represents the mean of the $ISE_{i,j,n}^{(k)}$, $k = 1, 2, \dots, M$, minus the lowest ISE mean for that line (i.e., minus the ISE mean of the best estimator for that specific target distribution and sample size). For each estimator ($j = 1, 2, \dots, 10$) and each sample size, the total of those differences to the best ISE means is calculated on the line called “total”. For each sample size, the lowest totals are highlighted in cyan. We see that Table 2 paints a nice picture of the asymmetric kernel c.d.f. estimators’ performance. Indeed, it shows that for each sample size ($n = 256, 1000$), the total of the differences to the best ISE means is significantly lower for the LogNormal (LN) and Birnbaum–Saunders (B–S) kernel c.d.f. estimators compared to all the other alternatives. For instance, the boundary kernel (BK) c.d.f. estimator would have been the go-to method in the past, but our results show that the total (over the eight target distributions) of the ISE mean differences to the best ISE means is more than three times lower for the LN and B–S kernel c.d.f. estimators compared to the BK c.d.f. estimator when $n = 256$, and similarly, it is more than two times lower for the LN and B–S kernel c.d.f. estimators compared to the BK c.d.f. estimator when $n = 1000$. Even if we put aside the best asymmetric kernel c.d.f. estimators, the totals of the ISE mean differences to the best ISE means for all the other asymmetric kernel c.d.f. estimators are also lower than for the BK c.d.f. estimator when $n = 256$, and they are in the same range (or better in the case of LN and B–S) when $n = 1000$. This means that all the asymmetric kernel estimators are overall better alternatives (or at least always remain competitive) compared to the BK estimator, although the advantage seems to dissipate (except for LN and B–S) when n increases.

10. Conclusions

In this paper, we considered five new asymmetric kernel c.d.f. estimators, namely the Gamma (Gam), inverse Gamma (IGam), LogNormal (LN), inverse Gaussian (IGau) and reciprocal inverse Gaussian (RIG) kernel c.d.f. estimators. We proved the asymptotic normality of these estimators and we also found asymptotic expressions for their bias, variance, mean squared error and mean integrated squared error. The expressions for

the optimal bandwidth under the mean integrated squared error were used in each case to implement a bandwidth selection procedure in our simulation study. With the same experimental design as Mombeni et al. [41] (but with an improved approximation of the integrals involved in the bandwidth selection procedures and the computation of the ISE's), our results show that the LogNormal and Birnbaum–Saunders kernel c.d.f. estimators perform the best overall. The results also show that all seven asymmetric kernel c.d.f. estimators are better in some cases and always at least competitive against the boundary kernel alternative presented by Tenreiro [57]. In that sense, all seven asymmetric kernel c.d.f. estimators are safe to use in place of more traditional methods. We recommend using the LogNormal and Birnbaum–Saunders kernel c.d.f. estimators in the future.

Author Contributions: Conceptualisation, methodology, investigation, formal analysis, software, coding and simulations, validation, visualisation, writing—original draft preparation, writing—review and editing, review of the literature, theoretical results and proofs, F.O.; software, coding and simulations, validation, P.L.d.M. All authors have read and agreed to the published version of the manuscript.

Funding: F. Ouimet is supported by postdoctoral fellowships from the NSERC (PDF) and the FRQNT (B3X supplement and B3XR).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The R code for the simulations in Section 8 is available online.

Acknowledgments: We thank Benedikt Funke for reminding us of the representation $\min\{T_1, T_2\} = \frac{1}{2}(T_1 + T_2) - \frac{1}{2}|T_1 - T_2|$, which helped tightening up the MSE and MISE results in Sections 6 and 7. This research includes computations using the computational cluster Katana supported by Research Technology Services at UNSW Sydney. We thank the referees for their insightful remarks which led to improvements in the presentation of this paper.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. Proof of the Results for the Gam Kernel

Proof of Lemma 1. If T denotes a random variable with the density

$$k_{\text{Gam}}(t | \alpha, \theta) = \frac{t^{\alpha-1} e^{-t/\theta}}{\theta^\alpha \Gamma(\alpha)}, \quad \text{with } (\alpha, \theta) = (b^{-1}x + 1, b), \quad (\text{A1})$$

then integration by parts yields

$$\begin{aligned} \mathbb{E}[\hat{F}_{n,b}^{\text{Gam}}(x)] - F(x) &= \mathbb{E}[F(T)] - F(x) \\ &= f(x) \cdot \mathbb{E}[T - x] + \frac{1}{2} f'(x) \cdot \mathbb{E}[(T - x)^2] + o_x(\mathbb{E}[(T - x)^2]) \\ &= f(x) \cdot b + \frac{1}{2} f'(x) \cdot b(2b + x) + o_x(b) \\ &= b \cdot (f(x) + \frac{x}{2} f'(x)) + o_x(b). \end{aligned} \quad (\text{A2})$$

Now, we want to compute the expression for the variance. Let S be a random variable with density $t \mapsto 2k_{\text{Gam}}(t | b^{-1}x + 1, b)\bar{K}_{\text{Gam}}(t | b^{-1}x + 1, b)$ and note that $\min\{T_1, T_2\}$ has that particular distribution if $T_1, T_2 \sim \text{Gam}(b^{-1}x + 1, b)$ are independent. Then, integration by parts and Corollary A1 yield, for any given $x \in (0, \infty)$,

$$\begin{aligned}
\mathbb{E}[\bar{K}_{\text{Gam}}^2(X_1, b^{-1}x + 1, b)] &= \mathbb{E}[F(S)] \\
&= F(x) + f(x) \cdot \mathbb{E}[S - x] + \mathcal{O}_x(\mathbb{E}[(S - x)^2]) \\
&= F(x) + f(x) \cdot \left[-\sqrt{\frac{bx}{\pi}} + \mathcal{O}_x(b)\right] + \mathcal{O}_x(b) \\
&= F(x) - b^{1/2} \cdot \frac{\sqrt{x}f(x)}{\sqrt{\pi}} + \mathcal{O}_x(b),
\end{aligned} \tag{A3}$$

so that

$$\begin{aligned}
\text{Var}(\hat{F}_{n,b}^{\text{Gam}}(x)) &= n^{-1} \mathbb{E}[\bar{K}_{\text{Gam}}^2(X_1, b^{-1}x + 1, b)] - n^{-1} (\mathbb{E}[\hat{F}_{n,b}^{\text{Gam}}(x)])^2 \\
&= n^{-1} F(x)(1 - F(x)) - n^{-1} b^{1/2} \cdot \frac{\sqrt{x}f(x)}{\sqrt{\pi}} + \mathcal{O}_x(n^{-1}b).
\end{aligned} \tag{A4}$$

This ends the proof. $\square$

Proof of Proposition 1. Note that $\hat{F}_{n,b}^{\text{Gam}}(x) - \mathbb{E}[\hat{F}_{n,b}^{\text{Gam}}(x)] = \frac{1}{n} \sum_{i=1}^n Z_{i,b}$, where

$$Z_{i,b} := \bar{K}_{\text{Gam}}(X_i | b^{-1}x + 1, b) - \mathbb{E}[\bar{K}_{\text{Gam}}(X_i | b^{-1}x + 1, b)], \quad 1 \leq i \leq n, \tag{A5}$$

are i.i.d. and centered random variables. It suffices to show the following Lindeberg condition for double arrays (see, e.g., Section 1.9.3 in [62]): for every $\varepsilon > 0$,

$$s_b^{-2} \mathbb{E}[Z_{1,b}^2 \mathbb{1}_{\{|Z_{1,b}| > \varepsilon n^{1/2} s_b\}}] \longrightarrow 0, \quad \text{as } n \rightarrow \infty, \tag{A6}$$

where $s_b^2 := \mathbb{E}[Z_{1,b}^2]$ and $b = b(n) \rightarrow 0$. This follows from the fact that $|Z_{1,b}| \leq 2$ for all $b > 0$, and $s_b = (n \text{Var}(\hat{F}_{n,b}^{\text{Gam}}))^{1/2} \rightarrow F(x)(1 - F(x))$ as $n \rightarrow \infty$ by Lemma 1. $\square$

Appendix B. Proof of the Results for the IGam Kernel

Proof of Lemma 2. If T denotes a random variable with the density

$$k_{\text{IGam}}(t | \alpha, \theta) = \frac{t^{-\alpha-1} e^{-1/(t\theta)}}{\theta^\alpha \Gamma(\alpha)}, \quad \text{with } (\alpha, \theta) = (b^{-1} + 1, x^{-1}b), \tag{A7}$$

then integration by parts yields (assuming $0 < b < 1/2$)

$$\begin{aligned}
\mathbb{E}[\hat{F}_{n,b}^{\text{IGam}}(x)] - F(x) &= \mathbb{E}[F(T)] - F(x) \\
&= f(x) \cdot \mathbb{E}[T - x] + \frac{1}{2} f'(x) \cdot \mathbb{E}[(T - x)^2] + o_x(\mathbb{E}[(T - x)^2]) \\
&= f(x) \cdot 0 + \frac{1}{2} f'(x) \cdot \frac{bx^2}{1-b} + o_x(b) \\
&= b \cdot \frac{x^2}{2} f'(x) + o_x(b).
\end{aligned} \tag{A8}$$

Now, we want to compute the expression for the variance. Let S be a random variable with density $t \mapsto 2k_{\text{IGam}}(t | b^{-1} + 1, x^{-1}b) \bar{K}_{\text{IGam}}(t | b^{-1} + 1, x^{-1}b)$ and note that $\min\{T_1, T_2\}$ has that particular distribution if $T_1, T_2 \sim \text{IGam}(b^{-1} + 1, x^{-1}b)$ are independent. Then, integration by parts and Corollary A2, for any given $x \in (0, \infty)$,

$$\begin{aligned}
\mathbb{E}[\bar{K}_{\text{IGam}}^2(X_1, b^{-1} + 1, x^{-1}b)] &= \mathbb{E}[F(S)] \\
&= F(x) + f(x) \cdot \mathbb{E}[S - x] + \mathcal{O}_x(\mathbb{E}[(S - x)^2]) \\
&= F(x) + f(x) \cdot \left[-x\sqrt{\frac{b}{\pi}} + \mathcal{O}_x(b)\right] + \mathcal{O}_x(b) \\
&= F(x) - b^{1/2} \cdot \frac{xf(x)}{\sqrt{\pi}} + \mathcal{O}_x(b), \tag{A9}
\end{aligned}$$

so that

$$\begin{aligned}
\mathbb{V}\text{ar}(\hat{F}_{n,b}^{\text{IGam}}(x)) &= n^{-1} \mathbb{E}[\bar{K}_{\text{IGam}}^2(X_1, b^{-1} + 1, x^{-1}b)] - n^{-1} (\mathbb{E}[\hat{F}_{n,b}^{\text{IGam}}(x)])^2 \\
&= n^{-1} F(x)(1 - F(x)) - n^{-1} b^{1/2} \cdot \frac{xf(x)}{\sqrt{\pi}} + \mathcal{O}_x(n^{-1}b). \tag{A10}
\end{aligned}$$

This ends the proof. $\square$

Proof of Proposition 3. Note that $\hat{F}_{n,b}^{\text{IGam}}(x) - \mathbb{E}[\hat{F}_{n,b}^{\text{IGam}}(x)] = \frac{1}{n} \sum_{i=1}^n Z_{i,b}$, where

$$Z_{i,b} := \bar{K}_{\text{IGam}}(X_i | b^{-1} + 1, x^{-1}b) - \mathbb{E}[\bar{K}_{\text{IGam}}(X_i | b^{-1} + 1, x^{-1}b)], \quad 1 \leq i \leq n, \tag{A11}$$

are i.i.d. and centered random variables. It suffices to show the following Lindeberg condition for double arrays (see, e.g., Section 1.9.3 in [62]): for every $\varepsilon > 0$,

$$s_b^{-2} \mathbb{E}[Z_{1,b}^2 \mathbb{1}_{\{|Z_{1,b}| > \varepsilon n^{1/2} s_b\}}] \longrightarrow 0, \quad \text{as } n \rightarrow \infty, \tag{A12}$$

where $s_b^2 := \mathbb{E}[Z_{1,b}^2]$ and $b = b(n) \rightarrow 0$. This follows from the fact that $|Z_{1,b}| \leq 2$ for all $b > 0$, and $s_b = (n \mathbb{V}\text{ar}(\hat{F}_{n,b}^{\text{IGam}}))^{1/2} \rightarrow F(x)(1 - F(x))$ as $n \rightarrow \infty$ by Lemma 2. $\square$

Appendix C. Proof of the Results for the LN Kernel

Proof of Lemma 3. If T denotes a random variable with the density

$$k_{\text{LN}}(t | \mu, \sigma) = \frac{1}{t\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(\log t - \mu)^2}{2\sigma^2}\right), \quad \text{with } (\mu, \sigma) = (\log x, \sqrt{b}), \tag{A13}$$

then integration by parts yields

$$\begin{aligned}
\mathbb{E}[\hat{F}_{n,b}^{\text{LN}}(x)] - F(x) &= \mathbb{E}[F(T)] - F(x) \\
&= f(x) \cdot \mathbb{E}[T - x] + \frac{1}{2} f'(x) \cdot \mathbb{E}[(T - x)^2] + \mathcal{O}_x(\mathbb{E}[(T - x)^2]) \\
&= f(x) \cdot x(e^{b/2} - 1) + \frac{1}{2} f'(x) \cdot x^2(e^{2b} - 2e^{b/2} + 1) + \mathcal{O}_x(b) \\
&= b \cdot \frac{x}{2} (f(x) + xf'(x)) + \mathcal{O}_x(b). \tag{A14}
\end{aligned}$$

Now, we want to compute the expression for the variance. Let S be a random variable with density $t \mapsto 2k_{\text{LN}}(t | \log x, \sqrt{b})\bar{K}_{\text{LN}}(t | \log x, \sqrt{b})$ and note that $\min\{T_1, T_2\}$ has that particular distribution if $T_1, T_2 \sim \text{LN}(\log x, \sqrt{b})$ are independent. Then, integration by parts and Corollary A3 yield, for any given $x \in (0, \infty)$,

$$\begin{aligned}
\mathbb{E}[\bar{K}_{\text{LN}}^2(X_1, \log x, \sqrt{b})] &= \mathbb{E}[F(S)] \\
&= F(x) + f(x) \cdot \mathbb{E}[S - x] + \mathcal{O}_x(\mathbb{E}[(S - x)^2]) \\
&= F(x) + f(x) \cdot [b^{1/2} \cdot \frac{-x}{\sqrt{\pi}} + \mathcal{O}_x(b)] + \mathcal{O}_x(b) \\
&= F(x) - b^{1/2} \cdot \frac{xf(x)}{\sqrt{\pi}} + \mathcal{O}_x(b),
\end{aligned} \tag{A15}$$

so that

$$\begin{aligned}
\text{Var}(\hat{F}_{n,b}^{\text{LN}}(x)) &= n^{-1} \mathbb{E}[\bar{K}_{\text{LN}}^2(X_1, \log x, \sqrt{b})] - n^{-1} (\mathbb{E}[\hat{F}_{n,b}^{\text{LN}}(x)])^2 \\
&= n^{-1} F(x)(1 - F(x)) - n^{-1} b^{1/2} \cdot \frac{xf(x)}{\sqrt{\pi}} + \mathcal{O}_x(n^{-1}b).
\end{aligned} \tag{A16}$$

This ends the proof. $\square$

Proof of Proposition 5. Note that $\hat{F}_{n,b}^{\text{LN}}(x) - \mathbb{E}[\hat{F}_{n,b}^{\text{LN}}(x)] = \frac{1}{n} \sum_{i=1}^n Z_{i,b}$, where

$$Z_{i,b} := \bar{K}_{\text{LN}}(X_i | \log x, \sqrt{b}) - \mathbb{E}[\bar{K}_{\text{LN}}(X_i | \log x, \sqrt{b})], \quad 1 \leq i \leq n, \tag{A17}$$

are i.i.d. and centered random variables. It suffices to show the following Lindeberg condition for double arrays (see, e.g., Section 1.9.3 in [62]): for every $\varepsilon > 0$,

$$s_b^{-2} \mathbb{E}[Z_{1,b}^2 \mathbb{1}_{\{|Z_{1,b}| > \varepsilon n^{1/2} s_b\}}] \longrightarrow 0, \quad \text{as } n \rightarrow \infty, \tag{A18}$$

where $s_b^2 := \mathbb{E}[Z_{1,b}^2]$ and $b = b(n) \rightarrow 0$. This follows from the fact that $|Z_{1,b}| \leq 2$ for all $b > 0$, and $s_b = (n \text{Var}(\hat{F}_{n,b}^{\text{LN}}))^{1/2} \rightarrow F(x)(1 - F(x))$ as $n \rightarrow \infty$ by Lemma 3. $\square$

Appendix D. Proof of the Results for the IGau Kernel

Proof of Lemma 4. If T denotes a random variable with the density

$$k_{\text{IGau}}(t | \mu, \lambda) = \sqrt{\frac{\lambda}{2\pi t^3}} \exp\left(-\lambda \frac{(t - \mu)^2}{2\mu^2 t}\right), \quad \text{with } (\mu, \lambda) = (x, b^{-1}x), \tag{A19}$$

then integration by parts yields

$$\begin{aligned}
\mathbb{E}[\hat{F}_{n,b}^{\text{IGau}}(x)] - F(x) &= \mathbb{E}[F(T)] - F(x) \\
&= f(x) \cdot \mathbb{E}[T - x] + \frac{1}{2} f'(x) \cdot \mathbb{E}[(T - x)^2] + o_x(\mathbb{E}[(T - x)^2]) \\
&= f(x) \cdot 0 + \frac{1}{2} f'(x) \cdot x^2 b + o_x(b) \\
&= b \cdot \frac{x^2}{2} f'(x) + o_x(b).
\end{aligned} \tag{A20}$$

Now, we want to compute the expression for the variance. Let S be a random variable with density $t \mapsto 2k_{\text{IGau}}(t | x, b^{-1}x) \bar{K}_{\text{IGau}}(t | x, b^{-1}x)$ and note that $\min\{T_1, T_2\}$, which can also be written as $\frac{1}{2}(T_1 + T_2) - \frac{1}{2}|T_1 - T_2|$, has that particular distribution if $T_1, T_2 \sim \text{IGau}(x, b^{-1}x)$ are independent. Then, integration by parts together with the fact that $\mathbb{E}[T_1] = \mathbb{E}[T_2] = x$ yield, for any given $x \in (0, \infty)$,

$$\begin{aligned}
\mathbb{E}[\bar{K}_{\text{IGau}}^2(X_1, x, b^{-1}x)] &= \mathbb{E}[F(S)] \\
&= F(x) + f(x) \cdot \mathbb{E}[S - x] + \mathcal{O}_x(\mathbb{E}[(S - x)^2]) \\
&= F(x) + f(x) \cdot \left\{ -\frac{1}{2} \mathbb{E}[|T_1 - T_2|] \right\} + \mathcal{O}_x(b) \\
&= F(x) - b^{1/2} \cdot \frac{f(x)}{2} \left[\lim_{b \rightarrow 0} b^{-1/2} \mathbb{E}[|T_1 - T_2|] \right] + \mathcal{O}_x(b), \quad (\text{A21})
\end{aligned}$$

so that

$$\begin{aligned}
\mathbb{V}\text{ar}(\hat{F}_{n,b}^{\text{IGau}}(x)) &= n^{-1} \mathbb{E}[\bar{K}_{\text{IGau}}^2(X_1, x, b^{-1}x)] - n^{-1} (\mathbb{E}[\hat{F}_{n,b}^{\text{IGau}}(x)])^2 \\
&= n^{-1} F(x)(1 - F(x)) \\
&\quad - n^{-1} b^{1/2} \cdot \frac{f(x)}{2} \left[\lim_{b \rightarrow 0} b^{-1/2} \mathbb{E}[|T_1 - T_2|] \right] + \mathcal{O}_x(n^{-1}b). \quad (\text{A22})
\end{aligned}$$

This ends the proof. $\square$

Proof of Proposition 7. Note that $\hat{F}_{n,b}^{\text{IGau}}(x) - \mathbb{E}[\hat{F}_{n,b}^{\text{IGau}}(x)] = \frac{1}{n} \sum_{i=1}^n Z_{i,b}$, where

$$Z_{i,b} := \bar{K}_{\text{IGau}}(X_i | x, b^{-1}x) - \mathbb{E}[\bar{K}_{\text{IGau}}(X_i | x, b^{-1}x)], \quad 1 \leq i \leq n, \quad (\text{A23})$$

are i.i.d. and centered random variables. It suffices to show the following Lindeberg condition for double arrays (see, e.g., Section 1.9.3 in [62]): for every $\varepsilon > 0$,

$$s_b^{-2} \mathbb{E}[Z_{1,b}^2 \mathbb{1}_{\{|Z_{1,b}| > \varepsilon n^{1/2} s_b\}}] \longrightarrow 0, \quad \text{as } n \rightarrow \infty, \quad (\text{A24})$$

where $s_b^2 := \mathbb{E}[Z_{1,b}^2]$ and $b = b(n) \rightarrow 0$. This follows from the fact that $|Z_{1,b}| \leq 2$ for all $b > 0$, and $s_b = (n \mathbb{V}\text{ar}(\hat{F}_{n,b}^{\text{IGau}}))^{1/2} \rightarrow F(x)(1 - F(x))$ as $n \rightarrow \infty$ by Lemma 4. $\square$

Appendix E. Proof of the Results for the RIG Kernel

Proof of Lemma 5. If T denotes a random variable with the density

$$k_{\text{RIG}}(t | \mu, \lambda) = \sqrt{\frac{\lambda}{2\pi t}} \exp\left(-\lambda \frac{(1 - \mu t)^2}{2\mu^2 t}\right), \quad \text{with } (\mu, \lambda) = (x^{-1}(1 - b)^{-1}, x^{-1}b^{-1}), \quad (\text{A25})$$

then integration by parts yields

$$\begin{aligned}
\mathbb{E}[\hat{F}_{n,b}^{\text{RIG}}(x)] - F(x) &= \mathbb{E}[F(T)] - F(x) \\
&= f(x) \cdot \mathbb{E}[T - x] + \frac{1}{2} f'(x) \cdot \mathbb{E}[(T - x)^2] + o_x(\mathbb{E}[(T - x)^2]) \\
&= f(x) \cdot 0 + \frac{1}{2} f'(x) \cdot x^2 b(1 + b) + o_x(b) \\
&= b \cdot \frac{1}{2} x^2 f'(x) + o_x(b). \quad (\text{A26})
\end{aligned}$$

Now, we want to compute the expression for the variance. Let S be a random variable with density $t \mapsto 2k_{\text{RIG}}(t | x^{-1}(1 - b)^{-1}, x^{-1}b^{-1})\bar{K}_{\text{RIG}}(t | x^{-1}(1 - b)^{-1}, x^{-1}b^{-1})$ and note that $\min\{T_1, T_2\}$, which can also be written as $\frac{1}{2}(T_1 + T_2) - \frac{1}{2}|T_1 - T_2|$, has that particular distribution if $T_1, T_2 \sim \text{RIG}(x^{-1}(1 - b)^{-1}, x^{-1}b^{-1})$ are independent. Then, integration by parts together with the fact that $\mathbb{E}[T_1] = \mathbb{E}[T_2] = x$ yield, for any given $x \in (0, \infty)$,

$$\begin{aligned}
\mathbb{E}[\bar{K}_{\text{RIG}}^2(X_1, x^{-1}(1-b)^{-1}, x^{-1}b^{-1})] &= \mathbb{E}[F(S)] \\
&= F(x) + f(x) \cdot \mathbb{E}[S - x] + \mathcal{O}_x(\mathbb{E}[(S - x)^2]) \\
&= F(x) + f(x) \cdot \left\{ -\frac{1}{2} \mathbb{E}[|T_1 - T_2|] \right\} + \mathcal{O}_x(b) \\
&= F(x) - b^{1/2} \cdot \frac{f(x)}{2} \left[\lim_{b \rightarrow 0} b^{-1/2} \mathbb{E}[|T_1 - T_2|] \right] + \mathcal{O}_x(b), \tag{A27}
\end{aligned}$$

so that

$$\begin{aligned}
\text{Var}(\hat{F}_{n,b}^{\text{RIG}}(x)) &= n^{-1} \mathbb{E}[\bar{K}_{\text{RIG}}^2(X_1, x^{-1}(1-b)^{-1}, x^{-1}b^{-1})] - n^{-1} (\mathbb{E}[\hat{F}_{n,b}^{\text{RIG}}(x)])^2 \\
&= n^{-1} F(x)(1 - F(x)) \\
&\quad - n^{-1} b^{1/2} \cdot \frac{f(x)}{2} \left[\lim_{b \rightarrow 0} b^{-1/2} \mathbb{E}[|T_1 - T_2|] \right] + \mathcal{O}_x(n^{-1}b). \tag{A28}
\end{aligned}$$

This ends the proof. $\square$

Proof of Proposition 9. Note that $\hat{F}_{n,b}^{\text{RIG}}(x) - \mathbb{E}[\hat{F}_{n,b}^{\text{RIG}}(x)] = \frac{1}{n} \sum_{i=1}^n Z_{i,b}$, where

$$\begin{aligned}
Z_{i,b} &:= \bar{K}_{\text{RIG}}(X_i | x^{-1}(1-b)^{-1}, x^{-1}b^{-1}) \\
&\quad - \mathbb{E}[\bar{K}_{\text{RIG}}(X_i | x^{-1}(1-b)^{-1}, x^{-1}b^{-1})], \quad 1 \leq i \leq n, \tag{A29}
\end{aligned}$$

are i.i.d. and centered random variables. It suffices to show the following Lindeberg condition for double arrays (see, e.g., Section 1.9.3 in [62]): for every $\varepsilon > 0$,

$$s_b^{-2} \mathbb{E}[Z_{1,b}^2 \mathbb{1}_{\{|Z_{1,b}| > \varepsilon n^{1/2} s_b\}}] \longrightarrow 0, \quad \text{as } n \rightarrow \infty, \tag{A30}$$

where $s_b^2 := \mathbb{E}[Z_{1,b}^2]$ and $b = b(n) \rightarrow 0$. This follows from the fact that $|Z_{1,b}| \leq 2$ for all $b > 0$, and $s_b = (n \text{Var}(\hat{F}_{n,b}^{\text{RIG}}))^{1/2} \rightarrow F(x)(1 - F(x))$ as $n \rightarrow \infty$ by Lemma 5. $\square$

Appendix F. Technical Lemmas

The lemma below computes the first two moments for the minimum of two i.i.d. random variables with a Gamma distribution. The proof is a slight generalization of the answer provided by Felix Marin in the following [MathStackExchange post](https://math.stackexchange.com/questions/3910094/how-to-compute-this-double-integral-involving-the-gamma-function) (<https://math.stackexchange.com/questions/3910094/how-to-compute-this-double-integral-involving-the-gamma-function>) (accessed on September 15th 2021).

Lemma A1. Let $X, Y \stackrel{\text{i.i.d.}}{\sim} \text{Gamma}(\alpha, \theta)$, then

$$\mathbb{E}[(\min\{X, Y\})^j] = \frac{\theta^j \Gamma(\alpha + j)}{\Gamma(\alpha)} - \frac{j\theta^j}{\sqrt{\pi}} \cdot \frac{\Gamma(\alpha + j - 1/2)}{\Gamma(\alpha)}, \quad j \in \{1, 2\}, \tag{A31}$$

where $\Gamma(\alpha) := \int_0^\infty t^{\alpha-1} e^{-t} dt$ denotes the gamma function. In particular, for all $x \in \mathbb{R}$,

$$\mathbb{E}[(\min\{X, Y\} - x)] = \theta\alpha - \frac{\theta}{\sqrt{\pi}} \cdot \frac{\Gamma(\alpha + 1/2)}{\Gamma(\alpha)} - x, \tag{A32}$$

$$\begin{aligned}
\mathbb{E}[(\min\{X, Y\} - x)^2] &= \theta^2 \alpha(\alpha + 1) - \frac{2\theta^2}{\sqrt{\pi}} \cdot \frac{\Gamma(\alpha + 3/2)}{\Gamma(\alpha)} \\
&\quad - 2x \left[\theta\alpha - \frac{\theta}{\sqrt{\pi}} \cdot \frac{\Gamma(\alpha + 1/2)}{\Gamma(\alpha)} \right] + x^2. \tag{A33}
\end{aligned}$$

Proof. Assume throughout the proof that $j \in \{1, 2\}$. By the simple change of variables $(u, v) = (x/\theta, y/\theta)$, we have

$$\begin{aligned}\mathbb{E}[(\min\{X, Y\})^j] &= 2 \int_0^\infty \int_0^\infty y^j \frac{(xy)^{\alpha-1} e^{-(x+y)/\theta}}{\theta^{2\alpha} \Gamma^2(\alpha)} \mathbb{1}_{[0, \infty)}(x-y) dx dy \\ &= \frac{2\theta^j}{\Gamma^2(\alpha)} \int_0^\infty \int_0^\infty u^{\alpha-1} e^{-u} v^{\alpha+j-1} e^{-v} \mathbb{1}_{[0, \infty)}(u-v) du dv.\end{aligned}\quad (\text{A34})$$

By the integral representation of the Heaviside function

$$\mathbb{1}_{[0, \infty)}(x) = \lim_{\varepsilon \rightarrow 0^+} \frac{1}{2\pi i} \int_{-\infty}^\infty \frac{1}{\tau - i\varepsilon} e^{ix\tau} d\tau, \quad (\text{A35})$$

the above is

$$\begin{aligned}&= \frac{2\theta^j}{\Gamma^2(\alpha)} \lim_{\varepsilon \rightarrow 0^+} \frac{1}{2\pi i} \int_{-\infty}^\infty \frac{1}{\tau - i\varepsilon} \underbrace{\int_0^\infty u^{\alpha-1} e^{-(1-i\tau)u} du}_{=(1-i\tau)^{-\alpha} \Gamma(\alpha)} \underbrace{\int_0^\infty v^{\alpha+j-1} e^{-(1+i\tau)v} dv}_{=(1+i\tau)^{-\alpha-j} \Gamma(\alpha+j)} d\tau \\ &= \frac{2\theta^j \Gamma(\alpha+j)}{\Gamma(\alpha)} \lim_{\varepsilon \rightarrow 0^+} \frac{1}{2\pi i} \int_{-\infty}^\infty \frac{(1+\tau^2)^{-\alpha} (1+i\tau)^{-j}}{\tau - i\varepsilon} d\tau \\ &= \frac{2\theta^j \Gamma(\alpha+j)}{\Gamma(\alpha)} \left\{ \text{P.V.} \frac{1}{2\pi i} \int_{-\infty}^\infty \frac{(1+\tau^2)^{-\alpha} (1+i\tau)^{-j}}{\tau} d\tau + \frac{1}{2} \right\},\end{aligned}\quad (\text{A36})$$

where δ denotes the Dirac delta function. The second term in the last brace is $1/2$ and the principal value is

$$\begin{aligned}&= \frac{1}{2\pi i} \int_0^\infty \frac{(1+\tau^2)^{-\alpha}}{\tau} \left[\frac{1}{(1+i\tau)^j} - \frac{1}{(1-i\tau)^j} \right] d\tau \\ &= -\frac{j}{\pi} \int_0^\infty (1+\tau^2)^{-\alpha-j} d\tau,\end{aligned}\quad (\text{A37})$$

where we crucially used the fact that $j \in \{1, 2\}$ to obtain the last equality. Putting all the work back in (A36), we obtain

$$\begin{aligned}\mathbb{E}[(\min\{X, Y\})^j] &= \frac{2\theta^j \Gamma(\alpha+j)}{\Gamma(\alpha)} \left\{ -\frac{j}{\pi} \int_0^\infty (1+\tau^2)^{-\alpha-j} d\tau + \frac{1}{2} \right\} \\ &= \frac{2\theta^j \Gamma(\alpha+j)}{\Gamma(\alpha)} \left\{ -\frac{j}{2\pi} \int_0^\infty t^{1/2-1} (1+t)^{-\alpha-j} dt + \frac{1}{2} \right\}.\end{aligned}\quad (\text{A38})$$

The remaining integral can be evaluated using Ramanujan's master theorem. Indeed, note that

$$\begin{aligned}(1+t)^{-\alpha-j} &= \sum_{k=0}^\infty \binom{-\alpha-j}{k} t^k = \sum_{k=0}^\infty \binom{\alpha+j+k-1}{k} (-t)^k \\ &= \sum_{k=0}^\infty \varphi(k) \frac{(-t)^k}{k!}, \quad \text{with } \varphi(z) := \frac{\Gamma(\alpha+j+z)}{\Gamma(\alpha+j)}.\end{aligned}\quad (\text{A39})$$

Therefore,

$$\int_0^\infty t^{1/2-1} (1+t)^{-\alpha-j} dt = \Gamma(1/2) \varphi(-1/2) = \frac{\sqrt{\pi} \Gamma(\alpha+j-1/2)}{\Gamma(\alpha+j)}.\quad (\text{A40})$$

By putting this result in (A38), we obtain

$$\mathbb{E}[(\min\{X, Y\})^j] = \frac{\theta^j \Gamma(\alpha + j)}{\Gamma(\alpha)} - \frac{j\theta^j}{\sqrt{\pi}} \cdot \frac{\Gamma(\alpha + j - 1/2)}{\Gamma(\alpha)}. \quad (\text{A41})$$

This ends the proof. $\square$

Corollary A1. Let $X, Y \stackrel{\text{i.i.d.}}{\sim} \text{Gamma}(b^{-1}x + 1, b)$ for some $x, b \in (0, \infty)$, then

$$\begin{aligned} \mathbb{E}[(\min\{X, Y\} - x)] &= b\left(\frac{x}{b} + 1\right) - \frac{b}{\sqrt{\pi}} \cdot \frac{\Gamma(\frac{x}{b} + 3/2)}{\Gamma(\frac{x}{b} + 1)} - x \\ &= -\sqrt{\frac{bx}{\pi}} + b + \mathcal{O}_x(b^{3/2}), \end{aligned} \quad (\text{A42})$$

$$\begin{aligned} \mathbb{E}[(\min\{X, Y\} - x)^2] &= b^2\left(\frac{x}{b} + 1\right)\left(\frac{x}{b} + 2\right) - \frac{2b^2}{\sqrt{\pi}} \cdot \frac{\Gamma(\frac{x}{b} + 5/2)}{\Gamma(\frac{x}{b} + 1)} \\ &\quad - 2x\left[x - \sqrt{\frac{bx}{\pi}} + b + \mathcal{O}_x(b^{3/2})\right] + x^2 \\ &= bx + \mathcal{O}_x(b^{3/2}). \end{aligned} \quad (\text{A43})$$

The lemma below computes the first two moments for the minimum of two i.i.d. random variables with an inverse Gamma distribution.

Lemma A2. Let $X, Y \stackrel{\text{i.i.d.}}{\sim} \text{InverseGamma}(\alpha, \theta)$ and assume $\alpha > 2$, then

$$\mathbb{E}[(\min\{X, Y\})^j] = \frac{\theta^{-j} \Gamma(\alpha - j)}{\Gamma(\alpha)} - \frac{j\theta^{-j}}{\sqrt{\pi}} \cdot \frac{\Gamma(\alpha - j) \Gamma(\alpha - 1/2)}{\Gamma^2(\alpha)}, \quad j \in \{1, 2\}, \quad (\text{A44})$$

where Φ denotes the c.d.f. of the standard normal distribution. In particular, for all $x \in \mathbb{R}$,

$$\mathbb{E}[(\min\{X, Y\} - x)] = \frac{\theta^{-1}}{\alpha - 1} \left[1 - \frac{1}{\sqrt{\pi}} \cdot \frac{\Gamma(\alpha - 1/2)}{\Gamma(\alpha)} \right] - x, \quad (\text{A45})$$

$$\begin{aligned} \mathbb{E}[(\min\{X, Y\} - x)^2] &= \frac{\theta^{-2}}{(\alpha - 1)(\alpha - 2)} \left[1 - \frac{2}{\sqrt{\pi}} \cdot \frac{\Gamma(\alpha - 1/2)}{\Gamma(\alpha)} \right] \\ &\quad - 2x \frac{\theta^{-1}}{\alpha - 1} \left[1 - \frac{1}{\sqrt{\pi}} \cdot \frac{\Gamma(\alpha - 1/2)}{\Gamma(\alpha)} \right] + x^2. \end{aligned} \quad (\text{A46})$$

Proof. Assume throughout the proof that $j \in \{1, 2\}$. By the simple change of variables $(u, v) = (x^{-1}/\theta, y^{-1}/\theta)$ and the reparametrization $\tilde{\alpha} := \alpha - j > 0$, we have

$$\begin{aligned} \mathbb{E}[(\min\{X, Y\})^j] &= 2 \int_0^\infty \int_0^\infty y^j \frac{(xy)^{-\alpha-1} e^{-(x^{-1}+y^{-1})/\theta}}{\theta^{2\alpha} \Gamma^2(\alpha)} \mathbb{1}_{[0, \infty)}(x - y) dx dy \\ &= \frac{2\theta^{-j}}{\Gamma^2(\alpha)} \int_0^\infty \int_0^\infty u^{\alpha-1} e^{-u} v^{\alpha-j-1} e^{-v} \mathbb{1}_{[0, \infty)}(v - u) du dv \\ &= \frac{2\theta^{-j}}{\Gamma^2(\alpha)} \int_0^\infty \int_0^\infty u^{\tilde{\alpha}+j-1} e^{-u} v^{\tilde{\alpha}-1} e^{-v} \mathbb{1}_{[0, \infty)}(v - u) du dv. \end{aligned} \quad (\text{A47})$$

We already evaluated this double integral in the proof of Lemma A1 (with α instead of $\tilde{\alpha}$). The above is

$$= \frac{2\theta^{-j}}{\Gamma^2(\alpha)} \cdot \frac{\Gamma(\tilde{\alpha})}{2} \left[\Gamma(\tilde{\alpha} + j) - \frac{j}{2\sqrt{\pi}} \cdot \Gamma(\tilde{\alpha} + j - 1/2) \right]. \quad (\text{A48})$$

This ends the proof. $\square$

Corollary A2. Let $X, Y \stackrel{\text{i.i.d.}}{\sim} \text{InverseGamma}(b^{-1} + 1, x^{-1}b)$ for some $x \in (0, \infty)$ and $b \in (0, 1)$, then

$$\begin{aligned}\mathbb{E}[(\min\{X, Y\} - x)] &= -\frac{x}{\sqrt{\pi}} \cdot \frac{\Gamma(b^{-1} + 1/2)}{\Gamma(b^{-1} + 1)} \\ &= -x\sqrt{\frac{b}{\pi}} + \mathcal{O}_x(b^{3/2}),\end{aligned}\quad (\text{A49})$$

$$\begin{aligned}\mathbb{E}[(\min\{X, Y\} - x)^2] &= x^2 \left[\frac{1}{1-b} - 1 \right] - \frac{2x^2}{\sqrt{\pi}} \frac{\Gamma(b^{-1} + 1/2)}{\Gamma(b^{-1} + 1)} \left[\frac{1}{1-b} - 1 \right] \\ &= bx^2 + \mathcal{O}_x(b^{3/2}).\end{aligned}\quad (\text{A50})$$

The lemma below computes the first two moments for the minimum of two i.i.d. random variables with a LogNormal distribution.

Lemma A3. Let $X, Y \stackrel{\text{i.i.d.}}{\sim} \text{LogNormal}(\mu, \sigma)$, then

$$\mathbb{E}[(\min\{X, Y\})^a] = 2e^{a\mu + \frac{(a\sigma)^2}{2}} \Phi\left(-\frac{a\sigma}{\sqrt{2}}\right), \quad a > 0, \quad (\text{A51})$$

where Φ denotes the c.d.f. of the standard normal distribution. In particular, for all $x \in \mathbb{R}$,

$$\begin{aligned}\mathbb{E}[(\min\{X, Y\} - x)] &= 2e^{\mu + \frac{\sigma^2}{2}} \Phi\left(-\frac{\sigma}{\sqrt{2}}\right) - x, \\ \mathbb{E}[(\min\{X, Y\} - x)^2] &= 2e^{2\mu + 2\sigma^2} \Phi\left(-\sqrt{2}\sigma\right) - 4xe^{\mu + \frac{\sigma^2}{2}} \Phi\left(-\frac{\sigma}{\sqrt{2}}\right) + x^2.\end{aligned}$$

Proof. With the change of variables

$$\begin{pmatrix} u \\ v \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}, \quad \left| \frac{d(u, v)}{d(x, y)} \right| = 1, \quad (\text{A52})$$

we have

$$\begin{aligned}\mathbb{E}[(\min\{X, Y\})^a] &= 2 \int_{-\infty}^{\infty} \int_y^{\infty} e^{a(\mu + \sigma y)} \frac{1}{2\pi} e^{-\frac{x^2 + y^2}{2}} dx dy \\ &= 2 \int_{-\infty}^{\infty} \int_0^{\infty} e^{a(\mu + \sigma \cdot \frac{-u+v}{\sqrt{2}})} \frac{1}{2\pi} e^{-\frac{u^2 + v^2}{2}} du dv \\ &= 2e^{a\mu + \frac{(a\sigma)^2}{2}} \int_{-\infty}^{\infty} \int_0^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{(u + \frac{a\sigma}{\sqrt{2}})^2}{2}} \frac{1}{\sqrt{2\pi}} e^{-\frac{(v - \frac{a\sigma}{\sqrt{2}})^2}{2}} du dv \\ &= 2e^{a\mu + \frac{(a\sigma)^2}{2}} \Phi\left(-\frac{a\sigma}{\sqrt{2}}\right).\end{aligned}\quad (\text{A53})$$

This ends the proof. $\square$

Corollary A3. Let $X, Y \stackrel{\text{i.i.d.}}{\sim} \text{LogNormal}(\log x, \sqrt{b})$ for some $x, b \in (0, \infty)$, then

$$\begin{aligned}\mathbb{E}[(\min\{X, Y\} - x)] &= x \left[2e^{b/2} \Phi\left(-\sqrt{\frac{b}{2}}\right) - 1 \right] \\ &= -x\sqrt{\frac{b}{\pi}} + \frac{bx}{2} + \mathcal{O}_x(b^{3/2}), \\ \mathbb{E}[(\min\{X, Y\} - x)^2] &= x^2 \left[2e^{2b} \Phi\left(-\sqrt{2b}\right) - 4e^{b/2} \Phi\left(-\sqrt{\frac{b}{2}}\right) + 1 \right] \\ &= bx^2 + \mathcal{O}_x(b^{3/2}).\end{aligned}$$

References

1. Aitchison, J.; Lauder, I.J. Kernel density estimation for compositional data. *J. Roy. Statist. Soc. Ser. C* **1985**, *34*, 129–137.
2. Chen, S.X. Beta kernel estimators for density functions. *Comput. Statist. Data Anal.* **1999**, *31*, 131–145.
3. Chen, S.X. Probability density function estimation using gamma kernels. *Ann. Inst. Statist. Math.* **2000**, *52*, 471–480.
4. Rosenblatt, M. Remarks on some nonparametric estimates of a density function. *Ann. Math. Statist.* **1956**, *27*, 832–837.
5. Parzen, E. On estimation of a probability density function and mode. *Ann. Math. Statist.* **1962**, *33*, 1065–1076.
6. Gasser, T.; Müller, H.G. Kernel estimation of regression functions. In *Smoothing Techniques for Curve Estimation*; Springer: Berlin/Heidelberg, Germany, 1979; pp. 23–68.
7. Rice, J. Boundary modification for kernel regression. *Comm. Statist. A—Theory Methods* **1984**, *13*, 893–900.
8. Gasser, T.; Müller, H.G.; Mammitzsch, V. Kernels for nonparametric curve estimation. *J. R. Statist. Soc. Ser. B* **1985**, *47*, 238–252.
9. Müller, H.G. Smooth optimum kernel estimators near endpoints. *Biometrika* **1991**, *78*, 521–530.
10. Zhang, S.; Karunamuni, R.J. On kernel density estimation near endpoints. *J. Statist. Plann. Inference* **1998**, *70*, 301–316.
11. Zhang, S.; Karunamuni, R.J. On nonparametric density estimation at the boundary. *J. Nonparametr. Statist.* **2000**, *12*, 197–221.
12. Bouezmarni, T.; Rolin, J.M. Consistency of the beta kernel density function estimator. *Canad. J. Statist.* **2003**, *31*, 89–98.
13. Renault, O.; Scaillet, O. On the way to recovery: A nonparametric bias free estimation of recovery rate densities. *J. Bank. Financ.* **2004**, *28*, 2915–2931.
14. Fernandes, M.; Monteiro, P.K. Central limit theorem for asymmetric kernel functionals. *Ann. Inst. Statist. Math.* **2005**, *57*, 425–442.
15. Hirukawa, M. Nonparametric multiplicative bias correction for kernel-type density estimation on the unit interval. *Comput. Statist. Data Anal.* **2010**, *54*, 473–495.
16. Bouezmarni, T.; Rombouts, J.V.K. Nonparametric density estimation for multivariate bounded data. *J. Statist. Plann. Inference* **2010**, *140*, 139–152.
17. Zhang, S.; Karunamuni, R.J. Boundary performance of the beta kernel estimators. *J. Nonparametr. Stat.* **2010**, *22*, 81–104.
18. Bertin, K.; Klutchnikoff, N. Minimax properties of beta kernel estimators. *J. Statist. Plann. Inference* **2011**, *141*, 2287–2297.
19. Bertin, K.; Klutchnikoff, N. Adaptive estimation of a density function using beta kernels. *ESAIM Probab. Stat.* **2014**, *18*, 400–417.
20. Igarashi, G. Bias reductions for beta kernel estimation. *J. Nonparametr. Stat.* **2016**, *28*, 1–30.
21. Jin, X.; Kawczak, J. Birnbaum-Saunders and lognormal kernel estimators for modelling durations in high frequency financial data. *Ann. Econ. Financ.* **2003**, *4*, 103–124. Available online: <http://aeconf.com/Articles/May2003/ae040106.pdf> (accessed on September 15th 2021).
22. Scaillet, O. Density estimation using inverse and reciprocal inverse Gaussian kernels. *J. Nonparametr. Stat.* **2004**, *16*, 217–226.
23. Bouezmarni, T.; Scaillet, O. Consistency of asymmetric kernel density estimators and smoothed histograms with application to income data. *Econom. Theor.* **2005**, *21*, 390–412.
24. Bouezmarni, T.; Rombouts, J.V.K. Density and hazard rate estimation for censored and α -mixing data using gamma kernels. *J. Nonparametr. Stat.* **2008**, *20*, 627–643.
25. Bouezmarni, T.; Rombouts, J.V.K. Nonparametric density estimation for positive time series. *Comput. Statist. Data Anal.* **2010**, *54*, 245–261.
26. Igarashi, G.; Kakizawa, Y. Re-formulation of inverse Gaussian, reciprocal inverse Gaussian, and Birnbaum-Saunders kernel estimators. *Statist. Probab. Lett.* **2014**, *84*, 235–246.
27. Igarashi, G.; Kakizawa, Y. Generalised gamma kernel density estimation for nonnegative data and its bias reduction. *J. Nonparametr. Stat.* **2018**, *30*, 598–639.
28. Charpentier, A.; Flachaire, E. Log-transform kernel density estimation of income distribution. *L'actualité économique Rev. D'analyse économique* **2015**, *91*, 141–159.
29. Igarashi, G. Weighted log-normal kernel density estimation. *Comm. Statist. Theory Methods* **2016**, *45*, 6670–6687.
30. Zougab, N.; Adjabi, S. Multiplicative bias correction for generalized Birnbaum-Saunders kernel density estimators and application to nonnegative heavy tailed data. *J. Korean Statist. Soc.* **2016**, *45*, 51–63.
31. Kakizawa, Y.; Igarashi, G. Inverse gamma kernel density estimation for nonnegative data. *J. Korean Statist. Soc.* **2017**, *46*, 194–207.
32. Kakizawa, Y. Nonparametric density estimation for nonnegative data, using symmetrical-based inverse and reciprocal inverse Gaussian kernels through dual transformation. *J. Statist. Plann. Inference* **2018**, *193*, 117–135.
33. Zougab, N.; Harfouche, L.; Ziane, Y.; Adjabi, S. Multivariate generalized Birnbaum-Saunders kernel density estimators. *Comm. Statist. Theory Methods* **2018**, *47*, 4534–4555.
34. Zhang, S. A note on the performance of the gamma kernel estimators at the boundary. *Statist. Probab. Lett.* **2010**, *80*, 548–557.
35. Kakizawa, Y. Multivariate non-central Birnbaum-Saunders kernel density estimator for nonnegative data. *J. Statist. Plann. Inference* **2020**, *209*, 187–207.
36. Ouimet, F.; Tolosana-Delgado, R. Asymptotic properties of Dirichlet kernel density estimators. *J. Multivar. Anal.* **2022**, *187*, 104832.
37. Kokonendji, C.C.; Libengué Dobélé-Kpoka, F.G.B. Asymptotic results for continuous associated kernel estimators of density functions. *Afr. Diaspora J. Math.* **2018**, *21*, 87–97.
38. Kokonendji, C.C.; Somé, S.M. On multivariate associated kernels to estimate general density functions. *J. Korean Statist. Soc.* **2018**, *47*, 112–126.
39. Kokonendji, C.C.; Somé, S.M. Bayesian bandwidths in semiparametric modelling for nonnegative orthant data with diagnostics. *Stats* **2021**, *4*, 162–183.

40. Hirukawa, M. *Asymmetric Kernel Smoothing*; SpringerBriefs in Statistics; Springer: Singapore, 2018; p. xii+110.
41. Mombeni, H.A.; Masouri, B.; Akhoond, M.R. Asymmetric Kernels for Boundary Modification in Distribution Function Estimation. *REVSTAT* **2019**, 1–27. Available online: <https://www.ine.pt/revstat/pdf/Asymmetrickernelsforboundarymodificationindistributionfunctionestimation.pdf> (accessed on September 15th 2021).
42. Babu, G.J.; Canty, A.J.; Chaubey, Y.P. Application of Bernstein polynomials for smooth estimation of a distribution and density function. *J. Statist. Plann. Inference* **2002**, 105, 377–392.
43. Leblanc, A. Chung-Smirnov property for Bernstein estimators of distribution functions. *J. Nonparametr. Stat.* **2009**, 21, 133–142.
44. Leblanc, A. On estimating distribution functions using Bernstein polynomials. *Ann. Inst. Statist. Math.* **2012**, 64, 919–943.
45. Leblanc, A. On the boundary properties of Bernstein polynomial estimators of density and distribution functions. *J. Statist. Plann. Inference* **2012**, 142, 2762–2778.
46. Dutta, S. Distribution function estimation via Bernstein polynomial of random degree. *Metrika* **2016**, 79, 239–263.
47. Jmaei, A.; Slaoui, Y.; Dellagi, W. Recursive distribution estimator defined by stochastic approximation method using Bernstein polynomials. *J. Nonparametr. Stat.* **2017**, 29, 792–805.
48. Erdoğan, M.S.; Dişibüyük, c.; Ege Oruç, O. An alternative distribution function estimation method using rational Bernstein polynomials. *J. Comput. Appl. Math.* **2019**, 353, 232–242.
49. Wang, X.; Song, L.; Sun, L.; Gao, H. Nonparametric estimation of the ROC curve based on the Bernstein polynomial. *J. Statist. Plann. Inference* **2019**, 203, 39–56.
50. Babu, G.J.; Chaubey, Y.P. Smooth estimation of a distribution and density function on a hypercube using Bernstein polynomials for dependent random vectors. *Statist. Probab. Lett.* **2006**, 76, 959–969.
51. Belalia, M. On the asymptotic properties of the Bernstein estimator of the multivariate distribution function. *Statist. Probab. Lett.* **2016**, 110, 249–256.
52. Dib, K.; Bouezmarni, T.; Belalia, M.; Kitouni, A. Nonparametric bivariate distribution estimation using Bernstein polynomials under right censoring. *Comm. Statist. Theory Methods* **2020**.
53. Ouimet, F. Asymptotic properties of Bernstein estimators on the simplex. *J. Multivariate Anal.* **2021**, 185, 104784.
54. Ouimet, F. On the boundary properties of Bernstein estimators on the simplex. *arXiv* **2021**, arXiv:2006.11756.
55. Hanebeck, A.; Klar, B. Smooth distribution function estimation for lifetime distributions using Szasz-Mirakyan operators. *Ann. Inst. Stat. Math.* **2021**.
56. Ouimet, F. On the Le Cam distance between Poisson and Gaussian experiments and the asymptotic properties of Szasz estimators. *J. Math. Anal. Appl.* **2021**, 499, 125033.
57. Tenreiro, C. Boundary kernels for distribution function estimation. *REVSTAT Stat. J.* **2013**, 11, 169–190.
58. Tiago de Oliveira, J. Estatística de densidades: Resultados assintóticos. *Rev. Fac. Ciências Lisb.* **1963**, 9, 111–206.
59. Nadaraja, E.A. Some new estimates for distribution functions. *Teor. Veroyatnost. i Primenen.* **1964**, 9, 550–554.
60. Watson, G.S.; Leadbetter, M.R. Hazard analysis. II. *Sankhyā Ser. A* **1964**, 26, 101–116.
61. Altman, N.; Léger, C. Bandwidth selection for kernel distribution function estimation. *J. Statist. Plann. Inference* **1995**, 46, 195–214.
62. Serfling, R.J. *Approximation Theorems of Mathematical Statistics*; Wiley Series in Probability and Mathematical Statistics; John Wiley & Sons, Inc.: New York, NY, USA, 1980; p. xvi+371.