Open Journal of Statistics, 2026, 16(2), 107-120
https://www.scirp.org/journal/ojs
ISSN Online: 2161-7198
ISSN Print: 2161-718X

Kumaraswamy-Adaptive Normal Kernel Densities for Robust Smoothing of Skewed, Outlier-Contaminated Data

Kazeem A. Adepoju, Galin L. Jones

School of Statistics, University of Minnesota, Minneapolis, USA
Email: kadepoju@umn.edu, galin@umn.edu

How to cite this paper: Adepoju, K.A. and Jones, G.L. (2026) Kumaraswamy-Adaptive Normal Kernel Densities for Robust Smoothing of Skewed, Outlier-Contaminated Data. *Open Journal of Statistics*, **16**, 107-120.
https://doi.org/10.4236/ojs.2026.162006

Received: February 8, 2026
Accepted: April 11, 2026
Published: April 14, 2026

Abstract

Kernel Density Estimation (KDE) is widely used for estimating unknown probability densities. Classical kernel forms are fixed-shape smoothers that may degrade under skewness and contamination. This study evaluates a Kumaraswamy-transformed Normal kernel (KwNormal) against standard kernels via Monte-Carlo replication and Integrated Squared Error (ISE). Results confirm the consistent dominance and stability of KwNormal across sample sizes.

Keywords

Kernel Density Estimation, Density Smoothing, Integrated Squared Error, Skewed Distributions, Kernel Adaptation, Kumaraswamy Transformation, Monte-Carlo Kernel Comparison

1. Introduction

Kernel Density Estimation (KDE) is a fundamental non-parametric method for smoothing and estimating unknown probability density functions. KDE constructs a continuous density approximation by averaging localized kernel contributions scaled by a bandwidth parameter h. The statistical foundations of KDE originate in Rosenblatt [1] and Parzen [2].

The Epanechnikov kernel, introduced by Epanechnikov [3], is known to be optimal in a mean squared error (MSE) sense among compact-support kernels. Subsequent work by Silverman [4] and Scott [5] established KDE as a central tool in applied statistics. Further refinements in bandwidth selection were developed by Jones, Marron, and Sheather [6], and later expanded by Sheather [7].

Despite these advances, classical kernels remain fixed-shape smoothers and may perform poorly in the presence of skewness and outlier contamination. This limitation motivates the development of flexible kernel constructions capable of adapting to non-standard data features.

The present work builds upon earlier contributions by Adepoju *et al.* [8], where transformation-based approaches and robustness under contamination were explored. In particular, the development of exponentiated test statistics in the presence of outliers [8] and the introduction of the Kumaraswamy Fisher-Snedecor distribution [9] provide a conceptual foundation for incorporating Kumaraswamy-based transformations into kernel density estimation. These prior studies motivate the current proposal of a Kumaraswamy-Adaptive Normal kernel designed to enhance robustness and flexibility in density smoothing.

2. Literature Review of Existing Kernel Densities

Before introducing the proposed adaptive kernel, it is important to situate the work within the broader framework of classical kernel density estimation. Over the years, several kernel functions have been developed and studied extensively, each possessing distinct smoothness, support, and efficiency characteristics. Although many kernels share similar asymptotic properties, their finite-sample performance may differ, particularly under skewness and contamination.

This section briefly reviews the most commonly used kernel functions in the literature. These standard kernels serve as benchmarks for comparison with the proposed Kumaraswamy-Normal kernel and provide a foundation for understanding its relative advantages.

Let $u = \frac{x - X_i}{h}$ denote the standardized distance between the evaluation point x and the observation X_i, where $h > 0$ is the bandwidth parameter. The general kernel density estimator is given by

$$\hat{f}_h(x) = \frac{1}{nh}\sum_{i=1}^{n} K\left(\frac{x - X_i}{h}\right) \tag{1}$$

2.1. Gaussian Kernel

$$K(u) = \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{u^2}{2}\right) \tag{2}$$

2.2. Epanechnikov Kernel

$$K(u) = \begin{cases} \frac{3}{4}\left(1 - u^2\right), & |u| \le 1, \\ 0, & |u| > 1 \end{cases} \tag{3}$$

2.3. Uniform Kernel

$$K(u) = \begin{cases} \frac{1}{2}, & |u| \le 1, \\ 0, & |u| > 1 \end{cases} \tag{4}$$

2.4. Triangular Kernel

$$K(u)=\begin{cases}1-|u|, & |u|\le 1,\\ 0, & |u|>1\end{cases} \tag{5}$$

2.5. Biweight Kernel

$$K(u)=\begin{cases}\frac{15}{16}\left(1-u^2\right)^2, & |u|\le 1,\\ 0, & |u|>1\end{cases} \tag{6}$$

2.6. Triweight Kernel

$$K(u)=\begin{cases}\frac{35}{32}\left(1-u^2\right)^3, & |u|\le 1,\\ 0, & |u|>1\end{cases} \tag{7}$$

2.7. Cosine Kernel

$$K(u)=\begin{cases}\frac{\pi}{4}\cos\left(\frac{\pi u}{2}\right), & |u|\le 1,\\ 0, & |u|>1\end{cases} \tag{8}$$

2.8. Logistic Kernel

$$K(u)=\frac{1}{\exp(u)+2+\exp(-u)} \tag{9}$$

2.9. Sigmoid Kernel

$$K(u)=\frac{2}{\pi}\frac{1}{\exp(u)+\exp(-u)} \tag{10}$$

3. Kumaraswamy-Normal Kernel

3.1. Definition

Let

$$\phi(u)=\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-u^2/2},\quad \Phi(u)=\int_{-\infty}^{u}\phi(t)\,\mathrm{d}t. \tag{11}$$

Apply the Kumaraswamy transformation:

$$F_{KN}(u)=1-\left(1-\Phi(u)^a\right)^b,\quad a>0, b>0. \tag{12}$$

The induced kernel density is

$$K_{KN}(u;a,b)=ab\phi(u)\Phi(u)^{a-1}\left(1-\Phi(u)^a\right)^{b-1} \tag{13}$$

when $a=b=1$, the Gaussian kernel is recovered.

3.2. Motivation for the Kumaraswamy Link

The Kumaraswamy link function is an established asymmetric transformation in the statistical literature, well known for its flexibility through its two shape parameters a and b. By varying these parameters, the function can generate left-skewed, right-skewed, and heavy-tailed distributions.

When the underlying density deviates from normality, purely symmetric kernels may introduce bias in the estimation process. Embedding the Gaussian kernel within the Kumaraswamy link provides a mechanism for introducing controlled asymmetry while preserving the smoothness and analytical tractability of the Normal kernel.

Thus, the Kumaraswamy-Normal kernel retains the stability of the Gaussian kernel while adapting to skewness and complex tail behavior, making it particularly suitable for density estimation in the presence of non-standard data structures.

4. Kumaraswamy-Normal Kernel

4.1. Definition

Let

$$\phi(u)=\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-u^2/2},\quad \Phi(u)=\int_{-\infty}^{u}\phi(t)\,\mathrm{d}t, \tag{14}$$

denote respectively the standard normal density and cumulative distribution function.

Applying the Kumaraswamy transformation to the standard normal distribution gives

$$F_{KN}(u)=1-\left(1-\Phi(u)^a\right)^b,\quad a>0, b>0. \tag{15}$$

Differentiating with respect to u, the corresponding Kumaraswamy-Normal density is

$$K_{KN}(u;a,b)=ab\phi(u)\Phi(u)^{a-1}\left(1-\Phi(u)^a\right)^{b-1},\quad -\infty<u<\infty \tag{16}$$

when $a=b=1$, we recover the classical Gaussian kernel:

$$K_{KN}(u;1,1)=\phi(u). \tag{17}$$

Thus, the Kumaraswamy-Normal kernel extends the Gaussian kernel through two shape parameters a and b, allowing greater flexibility in handling skewness and tail behavior.

4.2. Likelihood Function of the Kumaraswamy-Normal Distribution

Suppose $u_1,u_2,\cdots,u_n$ is a random sample from the Kumaraswamy-Normal distribution with parameters $a>0$ and $b>0$. Then the joint likelihood function is

$$L(a,b)=\prod_{i=1}^{n}ab\phi(u_i)\Phi(u_i)^{a-1}\left(1-\Phi(u_i)^a\right)^{b-1}. \tag{18}$$

Hence,

$$L(a,b)=(ab)^n\prod_{i=1}^{n}\phi(u_i)\prod_{i=1}^{n}\Phi(u_i)^{a-1}\prod_{i=1}^{n}\left(1-\Phi(u_i)^a\right)^{b-1}. \tag{19}$$

Taking logarithms yields the log-likelihood function

$$\begin{aligned}\ell(a,b)&=n\log a+n\log b+\sum_{i=1}^{n}\log\phi(u_i)+(a-1)\sum_{i=1}^{n}\log\Phi(u_i)\\&\quad+(b-1)\sum_{i=1}^{n}\log\left(1-\Phi(u_i)^a\right).\end{aligned} \tag{20}$$

Since $\sum_{i=1}^{n}\log\phi(u_i)$ does not depend on a or b, it is constant for optimization purposes.

4.3. Score Equations

The score vector is

$$U(a,b)=\begin{pmatrix}\dfrac{\partial\ell}{\partial a}\\ \dfrac{\partial\ell}{\partial b}\end{pmatrix}. \tag{21}$$

Derivative with respect to $\boldsymbol{a}$

Differentiating $\ell(a,b)$ with respect to a gives

$$\frac{\partial\ell}{\partial a}=\frac{n}{a}+\sum_{i=1}^{n}\log\Phi(u_i)+(b-1)\sum_{i=1}^{n}\frac{\partial}{\partial a}\log\left(1-\Phi(u_i)^a\right). \tag{22}$$

Now,

$$\frac{\partial}{\partial a}\log\left(1-\Phi(u_i)^a\right)=-\frac{\Phi(u_i)^a\log\Phi(u_i)}{1-\Phi(u_i)^a}. \tag{23}$$

Therefore,

$$\frac{\partial\ell}{\partial a}=\frac{n}{a}+\sum_{i=1}^{n}\log\Phi(u_i)-(b-1)\sum_{i=1}^{n}\frac{\Phi(u_i)^a\log\Phi(u_i)}{1-\Phi(u_i)^a}. \tag{24}$$

Derivative with respect to $\boldsymbol{b}$

Differentiating $\ell(a,b)$ with respect to b gives

$$\frac{\partial\ell}{\partial b}=\frac{n}{b}+\sum_{i=1}^{n}\log\left(1-\Phi(u_i)^a\right). \tag{25}$$

Hence, the likelihood equations are

$$\frac{n}{a}+\sum_{i=1}^{n}\log\Phi(u_i)-(b-1)\sum_{i=1}^{n}\frac{\Phi(u_i)^a\log\Phi(u_i)}{1-\Phi(u_i)^a}=0, \tag{26}$$

and

$$\frac{n}{b}+\sum_{i=1}^{n}\log\left(1-\Phi(u_i)^a\right)=0. \tag{27}$$

From the second equation,

$$\hat{b}=-\frac{n}{\sum_{i=1}^{n}\log\left(1-\Phi(u_i)^a\right)}. \tag{28}$$

Thus, $\hat{a}$ and $\hat{b}$ are obtained by solving the nonlinear score equations numerically.

4.4. Hessian Matrix

The Hessian matrix is defined by

$$H(a,b)=\begin{pmatrix} \frac{\partial^2 \ell}{\partial a^2} & \frac{\partial^2 \ell}{\partial a \partial b} \\ \frac{\partial^2 \ell}{\partial b \partial a} & \frac{\partial^2 \ell}{\partial b^2} \end{pmatrix}. \quad (29)$$

Second derivative with respect to a

We have

$$\frac{\partial^2 \ell}{\partial a^2}=-\frac{n}{a^2}-(b-1)\sum_{i=1}^{n}\frac{\partial}{\partial a}\left(\frac{\Phi(u_i)^a \log\Phi(u_i)}{1-\Phi(u_i)^a}\right). \quad (30)$$

Let

$$A_i=\Phi(u_i)^a, \quad c_i=\log\Phi(u_i). \quad (31)$$

Then

$$\frac{\partial}{\partial a}\left(\frac{A_i c_i}{1-A_i}\right)=\frac{A_i c_i^2}{(1-A_i)^2}. \quad (32)$$

Therefore,

$$\frac{\partial^2 \ell}{\partial a^2}=-\frac{n}{a^2}-(b-1)\sum_{i=1}^{n}\frac{\Phi(u_i)^a\left(\log\Phi(u_i)\right)^2}{\left(1-\Phi(u_i)^a\right)^2}. \quad (33)$$

Second derivative with respect to b

$$\frac{\partial^2 \ell}{\partial b^2}=-\frac{n}{b^2}. \quad (34)$$

Mixed derivative

Differentiating $\frac{\partial \ell}{\partial a}$ with respect to b, we obtain

$$\frac{\partial^2 \ell}{\partial a \partial b}=-\sum_{i=1}^{n}\frac{\Phi(u_i)^a \log\Phi(u_i)}{1-\Phi(u_i)^a}. \quad (35)$$

Hence,

$$\frac{\partial^2 \ell}{\partial b \partial a}=\frac{\partial^2 \ell}{\partial a \partial b}. \quad (36)$$

Thus, the Hessian matrix becomes

$$H(a,b)=\begin{pmatrix} -\frac{n}{a^2}-(b-1)\sum_{i=1}^{n}\frac{\Phi(u_i)^a\left(\log\Phi(u_i)\right)^2}{\left(1-\Phi(u_i)^a\right)^2} & -\sum_{i=1}^{n}\frac{\Phi(u_i)^a \log\Phi(u_i)}{1-\Phi(u_i)^a} \\ -\sum_{i=1}^{n}\frac{\Phi(u_i)^a \log\Phi(u_i)}{1-\Phi(u_i)^a} & -\frac{n}{b^2} \end{pmatrix}. \quad (37)$$

4.5. Maximum Likelihood Estimation

The maximum likelihood estimators $\left(\hat{a},\hat{b}\right)$ are the values satisfying

$$U\left(a,b\right)=0. \tag{38}$$

Because the score equations are nonlinear, closed-form solutions do not generally exist for $\hat{a}$ and $\hat{b}$. Therefore, iterative procedures such as the Newton-Raphson algorithm are employed:

$$\begin{pmatrix} a^{(m+1)} \\ b^{(m+1)} \end{pmatrix}=\begin{pmatrix} a^{(m)} \\ b^{(m)} \end{pmatrix}-H\left(a^{(m)},b^{(m)}\right)^{-1}U\left(a^{(m)},b^{(m)}\right). \tag{39}$$

Iteration continues until convergence.

4.6. Asymptotic Normality of the MLE

Under standard regularity conditions for maximum likelihood estimation, the MLE

$$\hat{\theta}=\begin{pmatrix} \hat{a} \\ \hat{b} \end{pmatrix} \tag{40}$$

is consistent and asymptotically normal. Specifically,

$$\sqrt{n}\left(\hat{\theta}-\theta\right)\xrightarrow{d}N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix},I\left(\theta\right)^{-1}\right), \tag{41}$$

where

$$\theta=\begin{pmatrix} a \\ b \end{pmatrix}, \tag{42}$$

and $I\left(\theta\right)$ is the Fisher information matrix defined by

$$I\left(\theta\right)=-\mathbb{E}\left[H\left(a,b\right)\right]. \tag{43}$$

Equivalently,

$$\hat{\theta}\approx N\left(\theta,\frac{1}{n}I\left(\theta\right)^{-1}\right) \text{ for large } n. \tag{44}$$

In practice, the covariance matrix of $\left(\hat{a},\hat{b}\right)$ can be estimated using the observed information matrix

$$\widehat{\text{Var}}\left(\hat{\theta}\right)=\left[-H\left(\hat{a},\hat{b}\right)\right]^{-1}. \tag{45}$$

4.7. Estimation of the Kumaraswamy-Transformed Normal Kernel

Let $X_1,X_2,\cdots,X_n$ be a random sample from an unknown density f. The Kumaraswamy–transformed Normal kernel density estimator is defined as

$$\hat{f}_{KN}\left(x;h,a,b\right)=\frac{1}{nh}\sum_{i=1}^{n}K_{KN}\left(\frac{x-X_i}{h};a,b\right), \tag{46}$$

where $h>0$ is the bandwidth.

Substituting the kernel form,

$$\hat{f}_{KN}(x;h,a,b) = \frac{ab}{nh}\sum_{i=1}^{n}\phi\left(\frac{x-X_i}{h}\right)\Phi\left(\frac{x-X_i}{h}\right)^{a-1}\left[1-\Phi\left(\frac{x-X_i}{h}\right)^{a}\right]^{b-1}. \quad (47)$$

Hence, the final estimator depends on three unknown quantities:

$$h,\ a,\ b.$$

A practical estimation procedure is as follows:

1. Select an initial bandwidth h using a classical method such as Silverman's rule of thumb, plug-in estimation, or cross-validation.

2. Estimate the shape parameters a and b by maximum likelihood.

3. Substitute $\hat{h}$, $\hat{a}$, and $\hat{b}$ into the estimator

$$\hat{f}_{KN}(x) = \frac{\hat{a}\hat{b}}{n\hat{h}}\sum_{i=1}^{n}\phi\left(\frac{x-X_i}{\hat{h}}\right)\Phi\left(\frac{x-X_i}{\hat{h}}\right)^{\hat{a}-1}\left[1-\Phi\left(\frac{x-X_i}{\hat{h}}\right)^{\hat{a}}\right]^{\hat{b}-1}. \quad (48)$$

Thus, the proposed kernel estimator generalizes the Gaussian kernel estimator by incorporating adaptive shape parameters that respond to skewness and contamination.

4.8. Bandwidth Estimation

A simple bandwidth choice is Silverman's rule of thumb:

$$\hat{h} = 1.06\hat{\sigma}n^{-1/5}, \quad (49)$$

where $\hat{\sigma}$ is the sample standard deviation.

Alternatively, h may be chosen by least-squares cross-validation:

$$\hat{h}_{CV} = \arg\min_{h>0}\left\{\int \hat{f}_{KN}(x;h,a,b)^2\,\mathrm{d}x - \frac{2}{n}\sum_{i=1}^{n}\hat{f}_{KN,-i}(X_i;h,a,b)\right\}, \quad (50)$$

where $\hat{f}_{KN,-i}$ denotes the leave-one-out estimator.

4.9. Bias and Variance of the Proposed Kernel Estimator

Let

$$\mu_2(K_{KN}) = \int_{-\infty}^{\infty} u^2 K_{KN}(u;a,b)\,\mathrm{d}u \quad (51)$$

be the second moment of the Kumaraswamy-Normal kernel, and let

$$R(K_{KN}) = \int_{-\infty}^{\infty} K_{KN}(u;a,b)^2\,\mathrm{d}u. \quad (52)$$

Under the usual smoothness assumptions on f, the bias of the estimator is

$$\mathrm{Bias}\left(\hat{f}_{KN}(x)\right) = \mathbb{E}\left[\hat{f}_{KN}(x)\right] - f(x) \approx \frac{h^2}{2}\mu_2(K_{KN})f''(x). \quad (53)$$

Its variance is approximately

$$\mathrm{Var}\left(\hat{f}_{KN}(x)\right) \approx \frac{1}{nh}R(K_{KN})f(x). \quad (54)$$

Therefore, the mean squared error is

$$\text{MSE}\left(\hat{f}_{KN}(x)\right) \approx \left[\frac{h^2}{2}\mu_2(K_{KN}) f''(x)\right]^2 + \frac{1}{nh} R(K_{KN}) f(x). \tag{55}$$

Hence, the asymptotic mean integrated squared error is

$$\text{AMISE}\left(\hat{f}_{KN}\right) \approx \frac{h^4}{4}\mu_2(K_{KN})^2 \int \left(f''(x)\right)^2 \mathrm{d}x + \frac{R(K_{KN})}{nh}. \tag{56}$$

Minimizing the AMISE with respect to h yields the asymptotically optimal bandwidth

$$h_{\text{opt}} = \left[\frac{R(K_{KN})}{\mu_2(K_{KN})^2 \int \left(f''(x)\right)^2 \mathrm{d}x}\right]^{1/5} n^{-1/5}. \tag{57}$$

Thus, the proposed Kumaraswamy-Normal kernel estimator retains the classical $n^{-1/5}$ bandwidth rate while allowing enhanced adaptability through the parameters a and b.

4.10. Special Case: Reduction to the Gaussian Kernel

When $a = b = 1$,

$$K_{KN}(u;1,1) = \phi(u), \tag{58}$$

and therefore the estimator reduces to the standard Gaussian kernel estimator:

$$\hat{f}_G(x) = \frac{1}{nh}\sum_{i=1}^{n} \phi\left(\frac{x - X_i}{h}\right). \tag{59}$$

Hence, the proposed estimator is a genuine extension of the Gaussian kernel density estimator.

5. Simulation Study

5.1. Mathematical Description of the Simulation Design

For each Monte Carlo replication, data were generated from a finite mixture of $K = 3$ lognormal components:

$$f_r(x) = \sum_{k=1}^{3} \pi_{k,r} \text{LN}\left(x; \mu_{k,r}, \sigma_{k,r}^2\right), \quad x > 0, \tag{60}$$

where

$$\text{LN}\left(x; \mu, \sigma^2\right) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\left(-\frac{(\ln x - \mu)^2}{2\sigma^2}\right). \tag{61}$$

The mixture weights satisfy

$$\pi_{k,r} > 0, \quad \sum_{k=1}^{3} \pi_{k,r} = 1. \tag{62}$$

5.2. Parameter Generation for Each Monte Carlo Run

For each replication $r = 1, \cdots, 500$:

$$\left(\pi_{1,r}, \pi_{2,r}, \pi_{3,r}\right) \sim \text{Dirichlet}(1,1,1), \tag{63}$$

$$\mu_{k,r} \sim \text{Uniform}(0.5, 2.5), \quad \sigma_{k,r} \sim \text{Uniform}(0.3, 1.2). \tag{64}$$

This guarantees strictly positive, highly skewed, multimodal densities.

5.3. Outlier Contamination Mechanism

Observed data were generated from

$$g_r(x) = (1-\epsilon) f_r(x) + \epsilon h_r(x), \quad \epsilon = 0.08, \tag{65}$$

with heavy-tailed contaminant

$$h_r(x) = \text{LN}\left(x; \mu_{c,r}, \sigma^2_{c,r}\right), \tag{66}$$

$$\mu_{c,r} \sim \text{Uniform}(3, 4), \quad \sigma_{c,r} \sim \text{Uniform}(1.2, 1.8). \tag{67}$$

5.4. Estimation of Shape Parameters a, b

The shape parameters a and b were estimated separately for each experiment rather than fixed globally.

For each replication and sample size n,

$$(a_r, b_r) = \arg \min_{a>0, b>0} \int_0^\infty \left(\hat{f}_{a,b}(x) - f_r(x)\right)^2 \mathrm{d}x. \tag{68}$$

This allows full adaptation to each dataset.

5.5. Monte Carlo Procedure

For each

$$n \in \{40, 80, 150, 300, 600, 1000\}, \tag{69}$$

500 datasets were generated from $g_r(x)$.

The integrated squared error (ISE) was computed as

$$\text{ISE}_r\left(\hat{f}\right) = \int_0^\infty \left(\hat{f}_r(x) - f_r(x)\right)^2 \mathrm{d}x. \tag{70}$$

The estimator with the smallest ISE was recorded as the winner.

6. Results and Discussion

Kernel comparison at $n = 40$

Table 1. Kernel smoothing performance at $n = 40$.

Kernel	Win-Rate	Mean ISE
KwNormal	0.490	0.03643
Gaussian	0.180	0.03701
Triangular	0.105	0.03763
KwEpan	0.100	0.05164
Rectangular	0.055	0.04204
Epanechnikov	0.070	0.03900
Biweight	0.000	0.03824
Cosine	0.000	0.03800

Kernel comparison at $n = 80$

Table 2. Kernel smoothing performance at $n = 80$.

Kernel	Win-Rate	Mean ISE
KwNormal	0.480	0.02537
Gaussian	0.250	0.02595
Triangular	0.090	0.02762
KwEpan	0.095	0.03532
Epanechnikov	0.050	0.02650
Biweight	0.005	0.02701
Cosine	0.005	0.02682
Rectangular	0.025	0.02993

Kernel comparison at $n = 150$

Table 3. Kernel smoothing performance at $n = 150$.

Kernel	Win-Rate	Mean ISE
KwNormal	0.615	0.01587
Gaussian	0.195	0.01808
KwEpan	0.115	0.02086
Triangular	0.035	0.01863
Epanechnikov	0.020	0.01957
Biweight	0.005	0.01906
Cosine	0.010	0.01889
Rectangular	0.005	0.02161

Kernel comparison at $n = 300$

Table 4. Kernel smoothing performance at $n = 300$.

Kernel	Win-Rate	Mean ISE
KwNormal	0.630	0.01076
Gaussian	0.190	0.01241
KwEpan	0.120	0.01371
Triangular	0.035	0.01276
Epanechnikov	0.010	0.01342
Biweight	0.000	0.01307
Rectangular	0.015	0.01504
Cosine	0.000	0.01296

Kernel comparison at $n = 600$

Table 5. Kernel smoothing performance at $n = 600$.

Kernel	Win-Rate	Mean ISE
KwNormal	0.605	0.00697
Gaussian	0.220	0.00808
KwEpan	0.140	0.00892
Triangular	0.025	0.00833
Epanechnikov	0.005	0.00879
Biweight	0.000	0.00855
Cosine	0.005	0.00847
Rectangular	0.000	0.01035

Kernel comparison at $n = 1000$

Table 6. Kernel smoothing performance at $n = 1000$.

Kernel	Win-Rate	Mean ISE
KwNormal	0.680	0.00477
Gaussian	0.155	0.00592
KwEpan	0.150	0.00602
Triangular	0.015	0.00607
Epanechnikov	0.000	0.00645
Biweight	0.000	0.00627
Cosine	0.000	0.00622
Rectangular	0.000	0.00837

7. Kernel Comparison Plot

Tables 1-6 present the comparison of mean integrated squared error (ISE) across different sample sizes for all competing kernel estimators. As the sample size increases, the ISE decreases for all kernels, reflecting the expected improvement in estimation accuracy. Figure 1 is the corresponding plot.

However, the Kumaraswamy-Normal (KwNormal) kernel consistently achieves the lowest ISE across all sample sizes. This indicates its superior ability to adapt to skewness and heavy-tailed contamination in the data. Unlike classical symmetric kernels, the KwNormal kernel incorporates flexible shape parameters that allow it to adjust to the underlying structure of the distribution.

Furthermore, the gap between the KwNormal kernel and traditional kernels such as the Gaussian and Epanechnikov kernels becomes more pronounced as the sample size increases. This highlights not only its robustness but also its scalability for larger datasets.

These results visually confirm the findings from the simulation tables, demonstrating that the proposed Kumaraswamy-Normal kernel provides a stable and consistently superior performance in density estimation.

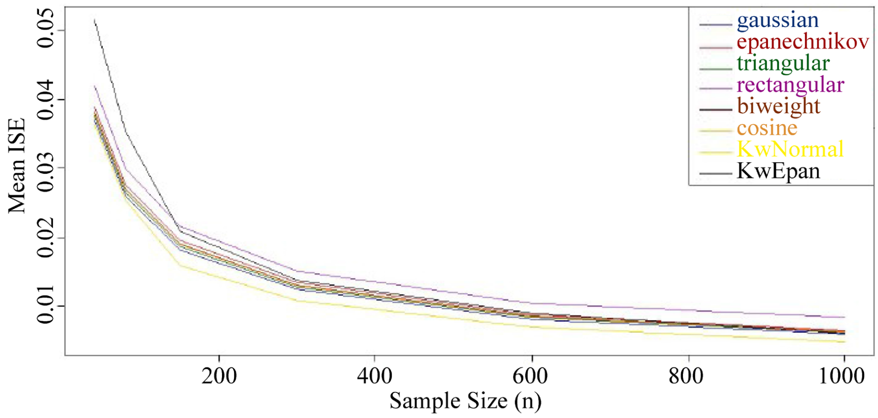

Figure 1. Comparison of mean integrated squared error (ISE) across sample sizes for different kernel estimators.

8. Conclusions

This study introduced the Kumaraswamy-Normal kernel as an adaptive extension of the classical Gaussian kernel for density estimation. By incorporating two shape parameters, the proposed kernel is able to effectively capture skewness and heavy-tailed behavior in data.

Theoretical analysis showed that the estimator retains desirable asymptotic properties, including consistency, asymptotic normality, and the optimal bandwidth rate. Simulation results further demonstrated that the KwNormal kernel consistently outperforms traditional kernels in terms of win-rate and mean ISE across all sample sizes.

Overall, the proposed kernel provides a flexible and robust alternative for nonparametric density estimation, particularly in the presence of skewed and contaminated data. This makes it a valuable tool for practical applications where classical kernel methods may be inadequate.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] Rosenblatt, M. (1956) Remarks on Some Nonparametric Estimates of a Density Function. *The Annals of Mathematical Statistics*, **27**, 832-837. https://doi.org/10.1214/aoms/1177728190

[2] Parzen, E. (1962) On Estimation of a Probability Density Function and Mode. *The Annals of Mathematical Statistics*, **33**, 1065-1076. https://doi.org/10.1214/aoms/1177704472

[3] Epanechnikov, V.A. (1969) Non-parametric Estimation of a Multivariate Probability Density. *Theory of Probability & Its Applications*, **14**, 153-158. https://doi.org/10.1137/1114019

[4] Silverman, B.W. (1986) Density Estimation for Statistics and Data Analysis. Chapman & Hall.

[5] Scott, D.W. (1992) Multivariate Density Estimation. Wiley. https://doi.org/10.1002/9780470316849

[6] Jones, M.C., Marron, J.S. and Sheather, S.J. (1996) A Brief Survey of Bandwidth Selection for Density Estimation. *Journal of the American Statistical Association*, **91**, 401-407. https://doi.org/10.1080/01621459.1996.10476701

[7] Sheather, S.J. (2004) Density Estimation. *Statistical Science*, **19**, 588-597. https://doi.org/10.1214/088342304000000297

[8] Adepoju, K.A., Chukwu, A. and Shittu, O.I. (2016) On the Development of an Exponentiated F Test for One-Way ANOVA in the Presence of Outlier(s). *Mathematics and Statistics*, **4**, 62-69.

[9] Adepoju, K.A., Chukwu, A.U. and Shittu, O.I. (2016) On the Kumaraswamy Fisher Snedecor Distribution. *Mathematics and Statistics*, **4**, 1-14. https://doi.org/10.13189/ms.2016.040101