On the Use of a Modified Intersection of Confidence Intervals (MICI_H) Kernel Density Estimation Approach

By Efosa Michael Ogbeide* & Joseph Erumnwosa Osemwenkhae±

Density estimation is an important aspect of statistics. Statistical inference often requires the knowledge of observed data density. A common method of density estimation is the kernel density estimation (KDE). It is a nonparametric estimation approach which requires a kernel function and a window size (smoothing parameter H). It aids density estimation and pattern recognition. So, this work focuses on the use of a modified intersection of confidence intervals (MICI_H) approach in estimating density. The Nigerian crime rate data reported to the Police as reported by the National Bureau of Statistics was used to demonstrate this new approach. This approach in the multivariate kernel density estimation is based on the data. The main way to improve density estimation is to obtain a reduced mean squared error (MSE), the errors for this approach was evaluated. Some improvements were seen. The aim is to achieve adaptive kernel density estimation. This was achieved under a sufficiently smoothing technique. This adaptive approach was based on the bandwidths selection. The quality of the estimates obtained of the MICI_H approach when applied, showed some improvements over the existing methods. The MICI_H approach has reduced mean squared error and relative faster rate of convergence compared to some other approaches. The approach of MICI_H has reduced points of discontinuities in the graphical densities the datasets. This will help to correct points of discontinuities and display adaptive density.

Keywords: approach, bandwidth, estimate, error, kernel density

Introduction

Data density estimation provides estimates of the probability function from which a set of data is drawn. Density is better estimated from the data. In density estimation, the true density is unknown. One of the popular approaches is the multivariate kernel density estimation. It is a nonparametric estimation approach which requires a kernel function and a bandwidth (window size or smoothing parameter H). Researches from Little and Rubin (2002) and Wu et al. (2007) showed that observation with missing data has a density curve with points of discontinuities that can be corrected when the missing data are accounted for in the original data set. This can be done according to Little and Rubin (2002), via good imputation method, which has comparatively lower mean squared error.

When we consider the variable window sizes on the multivariate cluster kernel density estimation (MCKDE) and the intersection of confidence interval (ICI)
approaches for estimating densities, we identified points for improvements, so that
the methods could be adaptive to the MKDE. In most cases, the above methods
could lead to under fitting, an indication that the methods are often less optimal
(Bowman and Azzalini 1997, Ogbeide et al. 2016). In this research work, we
propose data-driven approaches that require only the knowledge of the use of pilot
plots and the bandwidth sizes from the data set with a view to correcting the
identified problems, while aiming for lower asymptotic mean integrated squared
error (AMISE) and faster rates of convergence in the approaches. The aim of this
study is basically on how to fit density to multivariate data sets observations. That
is adaptive to the data at hand.

Researches have showed that that the performance of the kernel methods
depends largely on the smoothing parameter (window width) but depends very
little on the form of the kernel. According to Scott (1992) most times, analyses of
the multivariate data are more prevalent in practice than the univariate cases. The
Crucial problem in the multivariate kernel density estimation (MKDE) is to select
the window widths (bandwidth parameters) \( H \). The window widths control the
smoothness of the fitted density curve. The multivariate kernel density estimator
that we are going to study is a direct extension of the univariate estimator. Let
\( X_1, ..., X_n \) denote a \( d \)-variate random sample having a density \( f \). We shall use the
notation \( X_i = (X_{i1}, ..., X_{id})^T \) to denote the \( X_i \) and a generic vector \( x \in \mathbb{R}^d \) has
the representation \( x = (x_1, ..., x_d)^T \). The \( d \)-variate random sample \( X_1, ..., X_n \) drawn
from \( f \) the kernel estimator evaluated at \( x \) is given by;

\[
\hat{f}(x, H) = \frac{1}{n} \sum_{i=1}^{n} K_H(x - X_i)
\]

(1.1)

where \( n \) is the sample size, and \( H \) is a symmetric positive definite \( d \times d \) matrix
called the window widths, the smoothing parameters or the bandwidth matrix, and
\( K_H(x) = |H|^{-\frac{1}{2}} K(H^{-\frac{1}{2}} x) \), \(| . | \) stands for the determinant of \( H \) and \( K \) is \( d \)-variate
kernel satisfying \( \int k(x) dx = 1 \), where the integral is over \( \mathbb{R}^d \) unless stated
otherwise.

However, in choosing kernel to use, the gaussian kernel
\( K(u) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{u^2}{2}\right) \) is a popular choice among many kernels (Bowman and
Azzalini 1997, Katkovnik and Shmulevich 2002, Yang et al. 2019, Jayasinghe and
Jayasinghe 2021). The matrix \( H \) is a smoothing parameter and specifies the
‘width’ of the kernel around each sample point \( X_i \). The adaptive (smoothing)
methods are nonparametric density estimators that are sensitive to clustering/sparseness of sample values and other peculiarities, particularly at the
tails. Here the smoothing parameter \( H \) varies, hence the "adaptive" techniques.
This work is based on the analysis of a new modified intersection of confidence intervals (MICI_H). This approach is demonstrated using the Nigerian crime rate data reported to the Police as reported by the National Bureau of Statistics. This MICI_H would be compared to some other known approaches.

**Literature Review**

There are several research works considering the problem of window size selection in kernel density estimation (see Abramson 1982, Silverman 1986, Wand and Jones 1995, Bowman and Azzalini 1997, Katkovnik and Shmulevich 2002, Wu et al. 2007, Zhang and Chan 2011, Ogbeide et al. 2016, Tang et al. 2020). Considering the variable window sizes on the multivariate cluster sampling kernel density estimate (MMCKDE) and the intersection of confidence interval (ICI) rule, points for improvements are identified, so that the methods could be more adaptive. Generally, nonparametric densities are constructed with optimal window widths. There exist some methods of estimating bandwidths in the multivariate kernel density. Some of these methods use a fixed window width. However, the approach that uses varied window widths in the course of density estimation which seems adaptive are few (Wu et al. 2007, Ogbeide et al. 2017). A review of available variable methods showed basically that the cross-validation, the plug-in bandwidths approaches or any subjective method (which are fixed smoothing approaches) (see Duong and Hazelton 2005). There is the cluster and the average cluster approach by Wu and Tsai (2004) and Wu et al. (2007) which are more data sensitive are used. The window width controls the smoothness of the fitted density curve. The true density is unknown.

When we consider the studies on variable window sizes on the average cluster approach and the intersection of confidence interval (ICI) methods applied to MKDE, one is tasked with how sensitive these methods are, and the errors committed using these methods? What are the effects when we extend them to multivariate kernel density? These questions led to the reasons for their modifications. We identified points for improvements, so that the methods could be more adaptive (Ogbeide el al. 2017). Currently, a variety of sophistication of the basic kernel estimator has been proposed, all pointing to the importance of adaptive kernel estimator (see Kathovnik and Shmulevich 2002, Salgado-Ugarte and Perez-Hernandez 2003, Zhang and Chan 2011, Yang et al. 2019, Ogbeide and Osemwenkhae 2019). The “adaptive” nature of the density estimate arises from the varying bandwidth used in the estimation process. If \( h \), the bandwidth in (1.1) above, is “fixed” during estimation, we have the fixed kernel density estimation approach, but when it is allowed to vary all though the process of the estimation based on available data, we have the adaptive kernel method. A number of work considering the problem of kernel size selection exist (see Abramson 1982, Silverman 1986, Breiman et al. 1977, Hall 1990, Scott 1992, Wand and Jones 1995, Yang et al. 2019, Cortes and Sanz 2020, Ogbeide and Osemwenkhae 2021).
The most commonly used optimality criterion for selecting a bandwidth matrix is the mean integrated squared error (MISE)

\[ MISE(H) = E \left[ \int \frac{1}{h} \left( f(X) - f(X) \right)^2 dX \right] \tag{2.1} \]

where \( \int \) is a shorthand notation for \( \int_{R^n} \) and \( X \) is in \( n \) Euclidean plane \( R^n \). According to Horova et al. (2008), this equation (2.1) is in general does not have a closed-form expression, so we result to its asymptotic approximation (AMISE). Hence (2.1) could be factored as

\[ AMISE(H) \approx n^{-1} |H|^{-\frac{1}{2}} R(K) + \frac{1}{2} m_2(K) (vec^T H) \psi_4 (vec^T H) \tag{2.2} \]

where

- \( R(K) = \int K(X)^2 dX \), with \( R(K) = (4\pi)^{\frac{d}{2}} \) when \( K \) is a normal kernel.
- \( D^2 f \) is \( d \times d \) Hessian matrix of second order partial derivatives of \( f \).
- \( \psi_4 = \int (vecD^2 f(X))(vec^T D^2) dX \)
- \( D \) is a diagonal matrix with elements \( X_{11}, X_{22}, \ldots, X_{dd} \)
- \( vec \) is the vector operator which stacks the columns of a matrix into a single vector.

We observed that the quality of the AMISE to the MISE is given according to Horova et al. (2008) by

\[ MISE(H) = AMISE(H) + o(n^{-1} |H|^{-\frac{1}{2}} + trH^2) \tag{2.3} \]

where \( o \) indicates the usual \( o \) notation. This implies that AMISE is a ‘good’ approximation of the MISE as \( n \to \infty \). It has been shown that optimal bandwidth selector \( H \) has \( H = O(n^{-\frac{2}{d+4}}) \). According to Duong and Hazelton (2005) substituting this into equation (2.3) yields the optimal \( MISE(H) \) order as \( O(n^{-\frac{2}{d+4}}) \). The big O notation is applied element-wise. So when \( n \to \infty \), \( MISE \to 0 \). This implies the kernel density estimate converges in mean squared error and so also in probability to the true density \( f \). According to Wand and Jones (1995) and Horova et al. (2008), they asserted that it was better to estimate optimal MISE element-wise. They further asserted that the ideal optimal bandwidth selector that is point wise adaptive is given by
\[ H_{AMISE} = \arg \min_{h \in H} AMISE(H) \] (2.4)

Since this ideal bandwidth selector contains the unknown density function \( f \), that cannot be used directly. So some data density based approaches fixed the choice of bandwidth constant. However, we shall adopt point-wise adaptive bandwidth procedures in estimating densities.

The bandwidths used for the cluster approach by Wu et al. (2007) are optimal for information row/column (one dimensional) bandwidth per time in the multivariate data set. That is, it uses one bandwidth in the row or column during row/column cluster bandwidth selection. It is only row or column adaptive. Our approach is to make bandwidth selection to be data based on the smallest size of the row or column samples selections from the information matrix.

Clearly, in practice, one does not have access to the true density function \( f(x) \) which is to be estimated. Thus, a number of approaches can be taken for finding the bandwidth that will lead to better density estimation via varying the bandwidths (see Silverman 1986, Wand and Jones 1995, Katkovnik and Shmulevich 2002, Osemwenkhae and Ogbeide 2010). To this end, we modified the ICI approach in estimating density. The quality of the density estimates are assessed by comparing it to the density, obtained under the mean-squared error criterion. The error generated using these approach would be considered.

This work, present a data-driven methods that require the knowledge of pilot plot from optimal fixed window size and the variance of the estimate. This invariably reduces the amount of error at arriving at the “true density”. This work is basically concerned with an approach of achieving adaptive multivariate kernel density estimation. The aim of this study includes how to fit kernel density to the Nigerian Crime data reported to the Police from 2002-2006 with missing observations as reported by National Bureau of Statistics (2009) using the available data case (ADC) and mode-related expectation adaptive maximization (MEAM) approaches. See Ogbeide (2018) for details on MEAM imputation approach for dataset when missing observation occurs. Imputation is the way (best possible way) of correcting a dataset with missing observation when it occurs in a given dataset via statistical means in order not to affect the inference from the available dataset.

**Methodology**

In this section, the proposed method of estimating densities is presented. This method is the modified intersection of confidence intervals (MICI\(_H\)) approach. According to Wand and Jones (1995) and Horova et al. (2008), they asserted that it was better to estimate optimal MISE element-wise. They further asserted that the ideal optimal bandwidth selector that is point wise adaptive. This method is a modification of the ICI approach to density estimate. This modified approach adjusts the amount of bandwidths using some idea from the kernel nearest
neighbour estimation of the density to the multivariate data. Its smoothing parameter would be a \( n \times d \) dimensional matrix obtained from forming relevant number of clusters in an information matrix. The Euclidean distance would be used to form bandwidth.

The MICI\(_H\) procedure is basically a minimization of \( \text{AMISE}(H) \) with respect to \( H \), where it is equivalent to the selection of optimal \( h_j \) in \( \{H_1, H_2, \ldots, H_n\} \). Our data driven bandwidth matrix selector \( \hat{H} \) is point wise data base selection approach. Its density uses a pilot plot in order to address identified problem(s).

\[
\hat{H} = \text{agr} \min_{H \in \mathcal{H}} \text{AMISE}(H). \tag{3.1}
\]

Assuming that

\[
H = \{H_1 \leq H_2 \leq \ldots \leq H_n\} \tag{3.2}
\]
is a finite collection of window sizes, starting with a smallest \( h_j \in H \) and we determine a sequence of confidence intervals given by;

\[
D_{ij} = [L_{ij}, U_{ij}], \quad i=1,\ldots,n, \quad j=1,\ldots,d
\]

\[
\tilde{L}_{ij} = f(X_i) - \beta \text{std} \{f(X_i)\} \tag{3.3}
\]

\[
\tilde{U}_{ij} = f(X_i) + \beta \text{std} \{f(X_i)\}
\]
each \( h_j \) corresponding to a value in \( H_j \in H \). We assume the data at hand is normally distributed. Subjecting the data to normality, we propose \( \beta = 1.06 \) via normal reference rule.

And

\[
H_{\text{opt}}(X) = \left[ \frac{\text{abs} \tilde{L}_{oj}, U_{oj}}{v} \right] \tag{3.4}
\]

where \( \text{abs} \tilde{L}_{oj}, U_{oj} = \left| \tilde{L}_{oj} - U_{oj} \right| = \sqrt{\sum_{i=1}^{n} \sum_{j=1}^{d} \left( \tilde{L}_{oj} - U_{oj} \right)^2} \) (see Gray (1997) for lengths and distances’ details).

Subjectively we adopt \( v = 2 \), considering pilot plots. Where \( v \) is a positive real number. The \( \text{MICI}_H \) procedure is based on consideration of the intersection
of the adjusted intervals $D_{ij}$, $1 \leq i \leq n$ and $1 \leq j \leq d$. We adopt the bandwidth sizes $H_{opt,i}(X)$;

$$H_{opt}(X) = \left[ \frac{\text{abs}}{2} [L_{ij}, U_{ij}] \right] \text{ with } H_{opt,i}(X) \leq H_{opt,i+1}(X) \quad (3.5)$$

Consequently, substituting the bandwidths $H_{opt}(X)$ from equation (3.5) into the kernel density estimator

$$\hat{f}(X, H) = \frac{1}{n} \sum_{i=1}^{n} K_{h}(x - x_i)$$

to obtain the density estimates. The algorithm is as follows:

**Algorithm 1**

Step 1 $\bar{L} \leftarrow -\infty, U \leftarrow \infty$

Step 2 while $(L \leq U)$ and $(i \leq J)$ do

Step 3 $\tilde{L}_{ij} \leftarrow \hat{f}(X_i) - \beta \cdot \text{std} \{ \hat{f}(X_i) \}$

Step 4 $U_{ij} \leftarrow \hat{f}(X_i) + \beta \cdot \text{std} \{ \hat{f}(X_i) \}$

Step 5 $\bar{L}_{ij} \leftarrow \max[\bar{L}, \tilde{L}_{ij}], U_{j} \leftarrow \min[U, U_{ij}]$

Step 6 $i \leftarrow i + 1$

Step 7 $H_{opt,i}(x) = \left[ \frac{\text{abs}}{2} [\bar{L}_{ij}, U_{ij}] \right]$

Step 8 do $i \leftarrow i + 1$

Step 9 $H_{opt,i}(X) \leq H_{opt,i+1}(X)$

Step 10 compute $h_{ij}$ in $H_{ij} \in H$

Step 11 end while $(i = n)$.

**Application/Results**

The data for this study are drawn from the Nigerian National Bureau of Statistics Annual Abstract of Statistics published in the year 2009 on reported crime rate to the Police from 2002-2006 (www.nbs.gov.ng). The Tables of the density, error and convergence rate from the estimated bandwidths for the multivariate cluster sampling kernel density estimation (MCKDE) and the modified multivariate cluster sampling kernel density estimation (MMCKDE) and the MICI approaches from the Nigerian crime rate data are presented. Tables 1-3 show the data and the evaluations results calculated from the Nigerian crime rate data as reported to the police.
Table 1. Evaluated Estimates of Data with Missing Observation in Nigerian Crime Data Reported to the Police from 2002-2006 with Missing Observations as Reported by National Bureau of Statistics Annual Abstract of Statistics, 2009 Using the Available Data Case (ADC) and the Mode-Related Expectation Adaptive Maximization (MEAM) Approaches

<table>
<thead>
<tr>
<th>Offence</th>
<th>Case</th>
<th>2002</th>
<th>2003</th>
<th>2004</th>
<th>2005</th>
<th>2006</th>
</tr>
</thead>
<tbody>
<tr>
<td>False Pretence/Cheating</td>
<td>ADC</td>
<td>7913</td>
<td>9508</td>
<td>-</td>
<td>9580</td>
<td>6395</td>
</tr>
<tr>
<td></td>
<td>MEAM</td>
<td>7913</td>
<td>9508</td>
<td>9544</td>
<td>9580</td>
<td>6395</td>
</tr>
<tr>
<td>Unlawful Possession</td>
<td>ADC</td>
<td>3790</td>
<td>4142</td>
<td>5358</td>
<td>8772</td>
<td>8666</td>
</tr>
<tr>
<td></td>
<td>MEAM</td>
<td>1161</td>
<td>1289</td>
<td>2733</td>
<td>3892</td>
<td>7308</td>
</tr>
<tr>
<td>Receiving stolen property</td>
<td>ADC</td>
<td>199</td>
<td>148</td>
<td>-</td>
<td>631</td>
<td>473</td>
</tr>
<tr>
<td></td>
<td>MEAM</td>
<td>199</td>
<td>148</td>
<td>390</td>
<td>631</td>
<td>473</td>
</tr>
<tr>
<td>Arson</td>
<td>ADC</td>
<td>2005</td>
<td>1499</td>
<td>1289</td>
<td>1268</td>
<td>1010</td>
</tr>
<tr>
<td></td>
<td>MEAM</td>
<td>17</td>
<td>50</td>
<td>4</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>Perjury</td>
<td>ADC</td>
<td>7055</td>
<td>7298</td>
<td>7633</td>
<td>7967</td>
<td>5945</td>
</tr>
<tr>
<td></td>
<td>MEAM</td>
<td>7055</td>
<td>7298</td>
<td>7633</td>
<td>7967</td>
<td>5945</td>
</tr>
<tr>
<td>Breach of trust</td>
<td>ADC</td>
<td>220</td>
<td>272</td>
<td>-</td>
<td>99</td>
<td>132</td>
</tr>
<tr>
<td></td>
<td>MEAM</td>
<td>220</td>
<td>272</td>
<td>186</td>
<td>99</td>
<td>132</td>
</tr>
<tr>
<td>Escape from custody</td>
<td>ADC</td>
<td>2885</td>
<td>5171</td>
<td>-</td>
<td>3072</td>
<td>2610</td>
</tr>
<tr>
<td></td>
<td>MEAM</td>
<td>2885</td>
<td>5171</td>
<td>4122</td>
<td>3072</td>
<td>2610</td>
</tr>
<tr>
<td>Local acts</td>
<td>ADC</td>
<td>3262</td>
<td>3322</td>
<td>-</td>
<td>891</td>
<td>914</td>
</tr>
<tr>
<td></td>
<td>MEAM</td>
<td>3262</td>
<td>3322</td>
<td>2107</td>
<td>891</td>
<td>914</td>
</tr>
</tbody>
</table>

This result favours the use of the MEAM approach for missing data imputation. This is because, the MEAM have relative lower error propagation and relative faster convergence rates for this dataset (see Ogbeide 2018). An approach with these criteria of lower error propagation and faster convergence rates is preferred according to Little and Rubin (2002). Imputation is applied here obviously because the actual dataset with missing crime cases in the year not available could not be traced as reported. This is so, that we could have ‘clean’ data to operate and estimate density. Next we estimate the dataset densities from the data.

The densities estimates of the various bandwidths Nigerian minor crime rate data reported to the police from 2002-2006 with missing observations as reported by National Bureau of Statistics annual abstract of Statistics, 2009 with the MEAM imputation using the MCKDE, MMCKDE and MICI\(H\) approaches for false pretence/cheating, unlawful possession, received stolen goods, arson, perjury, gambling, breach of peace, escape from custody, local acts and other crimes from Table 2 shows that the densities sums are closer to unity in the MICI\(H\) than the other approaches. Furthermore, the bandwidth selections errors and the convergence rate from the estimated bandwidths for the multivariate cluster sampling kernel density estimation (MCKDE) and the modified multivariate cluster sampling kernel density estimation (MMCKDE) and the modified intersection of confidence intervals (MICI\(H\)) approaches from the Nigerian crime rate data reported for the false pretence/cheating, unlawful possession, received stolen goods, arson, perjury, gambling, breach of peace, escape from custody,
local acts and other crimes from Table 2 shows that the densities sums are closer to unity in the MCI than the other approaches The AMISE for MCI in false pretence/cheating, unlawful possession, received stolen goods, arson, perjury, gambling, breach of peace, escape from custody, local acts and other crimes were smaller than those of MMCKDE and MCKDE approaches.

| Table 2. Densities Estimates of the Various Bandwidths Nigerian Minor Crime Rate Data Reported to the Police from 2002-2006 with Missing Observations as Reported by National Bureau of Statistics Annual Abstract of Statistics, 2009 with the MEAM Imputation Using the MCKDE, MMCKDE and MCI Approaches |
|---|---|---|---|---|---|---|
| Offence | Case:Densities | 2002 | 2003 | 2004 | 2005 | 2006 | Density sum |
| False pretence/Cheating | Fixed H | 0.1842 | 0.2214 | 0.2231 | 0.2231 | 0.1411 | 0.9929 |
| | MCKDE | 0.1891 | 0.2123 | 0.2237 | 0.2231 | 0.1465 | 0.9947 |
| | MMCKDE | 0.1842 | 0.2189 | 0.2231 | 0.2231 | 0.1489 | 0.9982 |
| Unlawful possession | Fixed H | 0.1321 | 0.1352 | 0.1765 | 0.2991 | 0.1566 | 0.8995 |
| | MCKDE | 0.1321 | 0.1352 | 0.1765 | 0.2991 | 0.211 | 0.9539 |
| | MMCKDE | 0.1233 | 0.1331 | 0.1744 | 0.2577 | 0.282 | 0.9705 |
| Receiving | Fixed H | 0.0708 | 0.0787 | 0.1668 | 0.2399 | 0.3941 | 0.9503 |
| | MCKDE | 0.0811 | 0.0787 | 0.1668 | 0.2376 | 0.4023 | 0.9665 |
| | MMCKDE | 0.0811 | 0.0768 | 0.1881 | 0.2376 | 0.4147 | 0.9983 |
| Arson | Fixed H | 0.2514 | 0.1576 | 0.1322 | 0.1311 | 0.3195 | 0.9918 |
| | MCKDE | 0.2514 | 0.1713 | 0.1314 | 0.1231 | 0.3171 | 0.9943 |
| | MMCKDE | 0.2514 | 0.1713 | 0.1322 | 0.1244 | 0.3192 | 0.9985 |
| Perjury | Fixed H | 0.2412 | 0.1611 | 0.1371 | 0.1413 | 0.3192 | 0.9999 |
| | MCKDE | 0.2198 | 0.6088 | 0.0506 | 0.0379 | 0.0633 | 0.9804 |
| | MMCKDE | 0.2198 | 0.6198 | 0.0531 | 0.0399 | 0.0633 | 0.9959 |
| Gambling | Fixed H | 0.1081 | 0.0799 | 0.2001 | 0.3392 | 0.257 | 0.9843 |
| | MCKDE | 0.1083 | 0.0795 | 0.2116 | 0.3388 | 0.257 | 0.9952 |
| | MMCKDE | 0.1083 | 0.0801 | 0.2116 | 0.3428 | 0.257 | 0.9998 |
| Breach of Peace | Fixed H | 0.1965 | 0.2033 | 0.2102 | 0.2219 | 0.1656 | 0.9975 |
| | MCKDE | 0.1968 | 0.2001 | 0.2102 | 0.2219 | 0.1656 | 0.9946 |
| | MMCKDE | 0.1968 | 0.2033 | 0.2122 | 0.2219 | 0.1656 | 0.9998 |
| Escape from custody | Fixed H | 0.2422 | 0.2887 | 0.2041 | 0.1089 | 0.1453 | 0.9892 |
| | MCKDE | 0.2476 | 0.2887 | 0.2041 | 0.1089 | 0.1453 | 0.9946 |
| | MMCKDE | 0.2476 | 0.2939 | 0.2041 | 0.1089 | 0.1453 | 0.9998 |
| Local Acts | Fixed H | 0.1615 | 0.2833 | 0.2289 | 0.1701 | 0.1461 | 0.9899 |
| | MCKDE | 0.1615 | 0.2833 | 0.2298 | 0.1701 | 0.1461 | 0.9908 |
| | MMCKDE | 0.1615 | 0.2882 | 0.2308 | 0.172 | 0.1461 | 0.9986 |
| Others | Fixed H | 0.3108 | 0.2998 | 0.2007 | 0.0823 | 0.0871 | 0.9807 |
| | MCKDE | 0.3108 | 0.2998 | 0.2007 | 0.0849 | 0.0871 | 0.9833 |
| | MMCKDE | 0.3108 | 0.3123 | 0.2007 | 0.0849 | 0.0871 | 0.9958 |
| | MCI | 0.3108 | 0.3165 | 0.2007 | 0.0849 | 0.0871 | 1.0000 |
The $h^*$ (which is the error in relation to the fixed optimal bandwidth value), AMISE* and the convergence rates of methods are given below. Table 3 shows that there are reduced relative errors in the $h^*$ (which is the error in relation to the fixed optimal bandwidth value) and AMISE* in the proposed methods. The presented MICIM has faster convergence rates compared to the other versions. That is, the MICIM has lower error propagation and faster convergence rates when used to estimate the Nigerian crime rate data reported to the police. This is when compared to the fixed optimal H, MCKDE and the MMCKDE approaches. Generally, the AMISE shows the difference between the “true density” and the estimated density. The AMISE for MICIM is smaller than that of MMCKDE and MCKDE approaches.

Table 3. Bandwidth Selection Errors and the Convergence Rate from the Estimated Bandwidths for the Multivariate Cluster Sampling Kernel Density Estimation (MCKDE) and the Modified Multivariate Cluster Sampling Kernel Density Estimation (MMCKDE) and the MICIM Approaches from the Nigerian Crime Rate Data Reported

<table>
<thead>
<tr>
<th>Item(s)</th>
<th>Approach</th>
<th>Mean</th>
<th>Convergence Rate</th>
<th>Variance</th>
<th>$\hat{d}$</th>
<th>$h^*$</th>
<th>AMISE*</th>
</tr>
</thead>
<tbody>
<tr>
<td>False pretence/Cheating</td>
<td>MCKDE</td>
<td>24.450</td>
<td>1.0907</td>
<td>459.1561</td>
<td>21.4279</td>
<td>10.3869</td>
<td>2.1828x10^4</td>
</tr>
<tr>
<td></td>
<td>MMCKDE</td>
<td>17.450</td>
<td>1.6860</td>
<td>116.7649</td>
<td>10.8087</td>
<td>5.2384</td>
<td>1.2628x10^4</td>
</tr>
<tr>
<td></td>
<td>MICIM</td>
<td>20.153</td>
<td>1.4255</td>
<td>36.4999</td>
<td>6.0425</td>
<td>2.9285</td>
<td>7.9208x10^4</td>
</tr>
<tr>
<td>Unlawful possession</td>
<td>MCKDE</td>
<td>27.392</td>
<td>1.0000</td>
<td>559.0200</td>
<td>23.6436</td>
<td>11.4609</td>
<td>2.3616x10^1</td>
</tr>
<tr>
<td></td>
<td>MMCKDE</td>
<td>24.302</td>
<td>1.1991</td>
<td>179.3006</td>
<td>13.9030</td>
<td>6.4907</td>
<td>1.4985x10^1</td>
</tr>
<tr>
<td></td>
<td>MICIM</td>
<td>17.629</td>
<td>1.7740</td>
<td>37.6934</td>
<td>6.1334</td>
<td>2.9730</td>
<td>8.0243x10^4</td>
</tr>
<tr>
<td>Receiving stolen Property</td>
<td>MCKDE</td>
<td>33.146</td>
<td>1.1215</td>
<td>489.4700</td>
<td>22.1239</td>
<td>10.7242</td>
<td>2.2394x10^4</td>
</tr>
<tr>
<td></td>
<td>MMCKDE</td>
<td>30.580</td>
<td>1.1625</td>
<td>310.8864</td>
<td>17.6319</td>
<td>8.5468</td>
<td>1.8676x10^4</td>
</tr>
<tr>
<td></td>
<td>MICIM</td>
<td>30.259</td>
<td>1.1814</td>
<td>109.9777</td>
<td>10.4870</td>
<td>5.0834</td>
<td>1.2323x10^4</td>
</tr>
<tr>
<td>Arson</td>
<td>MCKDE</td>
<td>4.379</td>
<td>1.0285</td>
<td>10.3800</td>
<td>3.2218</td>
<td>1.5617</td>
<td>4.7943x10^1</td>
</tr>
<tr>
<td></td>
<td>MMCKDE</td>
<td>3.4910</td>
<td>1.3502</td>
<td>1.3664</td>
<td>1.1689</td>
<td>0.5661</td>
<td>2.1304x10^4</td>
</tr>
<tr>
<td></td>
<td>MICIM</td>
<td>3.5853</td>
<td>1.4018</td>
<td>0.1190</td>
<td>0.3449</td>
<td>0.1678</td>
<td>8.0242x10^4</td>
</tr>
<tr>
<td>Perjury</td>
<td>MCKDE</td>
<td>0.3340</td>
<td>1.0000</td>
<td>0.1400</td>
<td>0.3741</td>
<td>0.1813</td>
<td>8.5632x10^3</td>
</tr>
<tr>
<td></td>
<td>MMCKDE</td>
<td>0.2580</td>
<td>1.5881</td>
<td>0.0343</td>
<td>0.1852</td>
<td>0.0897</td>
<td>4.8793x10^3</td>
</tr>
<tr>
<td></td>
<td>MICIM</td>
<td>0.2474</td>
<td>1.5176</td>
<td>0.0177</td>
<td>0.1330</td>
<td>0.0644</td>
<td>3.743x10^3</td>
</tr>
<tr>
<td>Gambling</td>
<td>MCKDE</td>
<td>3.0290</td>
<td>0.5795</td>
<td>1.1800</td>
<td>1.0862</td>
<td>0.5265</td>
<td>2.0089x10^4</td>
</tr>
<tr>
<td></td>
<td>MMCKDE</td>
<td>3.0209</td>
<td>0.5829</td>
<td>1.0419</td>
<td>1.0070</td>
<td>0.4947</td>
<td>1.9914x10^4</td>
</tr>
<tr>
<td></td>
<td>MICIM</td>
<td>1.5082</td>
<td>1.7087</td>
<td>0.1698</td>
<td>0.4120</td>
<td>0.1997</td>
<td>9.2500x10^3</td>
</tr>
<tr>
<td>Breach of Peace</td>
<td>MCKDE</td>
<td>14.7749</td>
<td>1.058y</td>
<td>164.1300</td>
<td>12.8113</td>
<td>6.2101</td>
<td>1.4465x10^5</td>
</tr>
<tr>
<td></td>
<td>MMCKDE</td>
<td>12.0539</td>
<td>1.4273</td>
<td>72.6648</td>
<td>8.5243</td>
<td>4.1320</td>
<td>1.0441x10^4</td>
</tr>
<tr>
<td></td>
<td>MICIM</td>
<td>10.9229</td>
<td>1.6005</td>
<td>28.8064</td>
<td>5.371</td>
<td>2.6016</td>
<td>7.2117x10^4</td>
</tr>
<tr>
<td>Escape from custody</td>
<td>MCKDE</td>
<td>1.1721</td>
<td>1.1206</td>
<td>0.1000</td>
<td>0.3162</td>
<td>0.1532</td>
<td>7.4854x10^3</td>
</tr>
<tr>
<td></td>
<td>MMCKDE</td>
<td>1.1602</td>
<td>1.1379</td>
<td>0.0432</td>
<td>0.2078</td>
<td>0.1007</td>
<td>5.5851x10^3</td>
</tr>
<tr>
<td></td>
<td>MICIM</td>
<td>0.7484</td>
<td>1.9233</td>
<td>0.5404</td>
<td>0.2324</td>
<td>0.1126</td>
<td>5.3501x10^3</td>
</tr>
<tr>
<td>Local acts</td>
<td>MCKDE</td>
<td>22.5310</td>
<td>1.0000</td>
<td>124.0000</td>
<td>11.5711</td>
<td>5.6118</td>
<td>1.3339x10^4</td>
</tr>
<tr>
<td></td>
<td>MMCKDE</td>
<td>18.6925</td>
<td>1.3152</td>
<td>28.3639</td>
<td>5.3447</td>
<td>2.5907</td>
<td>7.1876x10^4</td>
</tr>
<tr>
<td></td>
<td>MICIM</td>
<td>15.3892</td>
<td>1.6653</td>
<td>28.3639</td>
<td>5.3257</td>
<td>2.5815</td>
<td>7.1671x10^4</td>
</tr>
<tr>
<td>Others</td>
<td>MCKDE</td>
<td>14.7482</td>
<td>0.7516</td>
<td>78.3682</td>
<td>8.8525</td>
<td>4.2911</td>
<td>1.0762x10^4</td>
</tr>
<tr>
<td></td>
<td>MMCKDE</td>
<td>10.7482</td>
<td>1.2189</td>
<td>62.2800</td>
<td>7.8919</td>
<td>3.8254</td>
<td>9.8171x10^3</td>
</tr>
<tr>
<td></td>
<td>MICIM</td>
<td>10.6617</td>
<td>1.2665</td>
<td>33.3262</td>
<td>5.7728</td>
<td>2.7982</td>
<td>7.6446x10^3</td>
</tr>
</tbody>
</table>

The estimated bandwidth selection errors and the convergence rates from the Nigerian crime rate data set with missing observations, via the various methods favour the use of the MICIM approach over the other approaches. This is because its bandwidth errors are smaller as well as having higher convergence rate. The
MMCKDE has some improvement over the MCKDE approach. These can be seen in Tables 2-3. Generally, the AMISE shows the difference between the “true density” and the estimated density. The AMISE for MICI\(_H\) is smaller than that of MMCKDE and MCKDE approaches. Graphical densities displays are presented next.

**False Pretence/Cheating**

**Figure 1a.** Graphical Density Estimates for False Pretence/Cheating Using the Fixed H Approach

**Figure 1b.** Graphical Density Estimates for False Pretence/Cheating Using the MCKDE Approach

**Figure 1c.** Graphical Density Estimates for False Pretence/Cheating Using the MMCKDE Approach

**Figure 1d.** Graphical Density Estimates for False Pretence/Cheating Using the MICI\(_H\) Approach
Unlawful Possession

**Figure 2a.** Graphical Density Estimates for Unlawful Possession Using the Fixed H Approach

**Figure 2b.** Graphical Density Estimates for Unlawful Possession Using the MCKDE Approach

**Figure 2c.** Density Graphical Density Estimates for Unlawful Possession Using the MMCKDE Approach

**Figure 2d.** Graphical Density Estimates for Unlawful Possession Using the MICI\(_H\) Approach
Receiving Stolen Properties

**Figure 3a.** Graphical Density Estimates for Receiving Stolen Property Using the Fixed H Approach

**Figure 3b.** Graphical Density Estimates for Receiving Stolen Property Using the MCKDE Approach

**Figure 3c.** Graphical Density Estimates for Receiving Stolen Property Using the MMCKDE Approach

**Figure 3d.** Graphical Density Estimates Receiving Stolen Property Using the MICI Approach
Arson

**Figure 4a.** Graphical Density Estimates for Arson Using the Fixed H Approach

**Figure 4b.** Graphical Density Estimates for Arson Using the MCKDE Approach

**Figure 4c.** Graphical Density Estimates for Arson Using the MMCKDE Approach

**Figure 4d.** Graphical Density Estimates for Arson Using the MICI Approach
Perjury

Figure 5a. Graphical Density Estimates for Perjury Using the Fixed H Approach

Figure 5b. Graphical Density Estimates for Perjury Using the MCKDE Approach

Figure 5c. Graphical Density Estimates for Perjury Using the MMCKDE Approach

Figure 5d. Graphical Density Estimates for Perjury Using the MICI Approach
Gambling

**Figure 6a.** Graphical Density Estimates for Gambling Using the Fixed H Approach

**Figure 6b.** Graphical Density Estimates for Gambling Using the MCKDE Approach

**Figure 6c.** Graphical Density Estimates for Gambling Using the MMCKDE Approach

**Figure 6d.** Graphical Density Estimates for Gambling Using the MIC₁ Approach
Breach of Peace

Figure 7a. Graphical Density Estimates for Breach of Peace Using the MCKDE Approach

Figure 7b. Graphical Density Estimates for Breach of Peace Using the Fixed H Approach

Figure 7c. Graphical Density Estimates for Breach of Peace Using the MMCKDE Approach

Figure 7d. Graphical Density Estimates for Breach of Peace Using the MICI_H Approach
Escape from Custody

**Figure 8a.** Graphical Density Estimates for Escape from Custody Using the fixed H Approach

**Figure 8b.** Graphical Density Estimates for Escape from Custody Using the MCKDE Approach

**Figure 8c.** Graphical Density Estimates for Escape from Custody Using the MMCKDE Approach

**Figure 8d.** Graphical Density Estimates for Escape from Custody Using the MICI Approach
Local Acts Crime

**Figure 9a.** Graphical Density Estimates for Local Acts Crime Using the fixed H Approach

**Figure 9b.** Graphical Density Estimates for Local Acts Crime Using the MCKDE Approach

**Figure 9c.** Graphical Density Estimates for Local Acts Crime Using the MMCKDE Approach

**Figure 9d.** Graphical Density Estimates for Local Acts Crime Using the MICI Approach
Other Crimes

The various approaches have identifiable differences from Figures 1a-10d, using the fixed H, MCKDE, MMCKDE and MICI\(H\) for the dataset in Nigerian crime rate data reported by the Nigerian National Bureau of Statistics. There are visible signs of discontinuities in the fixed H and the MCKDE graphical densities estimates of the data sets. These are seen in Figures 1a-10d as well as in Tables 2-3. The proposed approach of MICI\(H\) has reduced points of discontinuities in the graphical densities of the datasets compared to the fixed H, MCKDE, MMCKDE approaches. This MICI\(H\) approach helps to correct points of discontinuities and display adaptive density. The density from the fixed H bandwidth could serve as a pilot guide. The density from the fixed H bandwidth is by no means as adaptive as the variable bandwidths approaches.
Figure 11. Density for False Pretence/Cheating Using the Fixed H, MCKDE, MMCKDE and MIC_I Approaches

Figure 12. Density for Unlawful Possession Using the Fixed H, MCKDE, MMCKDE and MIC_I Approaches

Figure 13. Density for Receiving Stolen Property Using the Fixed H, MCKDE, MMCKDE and MIC_I Approaches

Figure 14. Density for Arson Using the Fixed H, MCKDE, MMCKDE and MIC_I Approaches

Figure 15. Density for Perjury Using the Fixed H, MCKDE, MMCKDE and MIC_I Approaches

Figure 16. Density for Gambling Using the Fixed H, MCKDE, MMCKDE and MIC_I Approaches
The various approaches have identifiable differences from Figures 11-20, using the fixed H, MCKDE, MMCKDE and MICI\textsubscript{H} for the Nigerian crime rate data reported.

The MICI\textsubscript{H} scheme produces better density estimation when compared to the fixed bandwidths H, MCKDE and MMCKDE approaches seeing the figures. This is a good approach because it is data sensitive and does not unfit or over fit the density curves as the case may be. This contributes significantly to the density estimates by showing more hidden features of the density. This approach provides and uses full bandwidths matrix from the data for the multivariate kernel density.
estimation. This approach has additional procedures in the choice and application of the smoothing parameters (bandwidths $H$) to multivariate density estimation. MICI$_H$ generate full bandwidths matrices. The cost of these steps brings about the adaptive density to be constructed. A study for less time efficient approach should be the aim of further research that vary the bandwidths selection based on the data in kernel density estimation.

The MICI$_H$ approach in estimating density is presented. The quality of the density estimates was assessed; comparing it with the densities obtained using the mean-squared error criterion showed some improvements. The crime data distribution/density is always needed for crime prevention and comparative analysis from time to time. The MICI$_H$ will be of practical help in discovery of the real dataset density.

Conclusions

This work presents a modified intersection of confidence intervals (MICI$_H$) approach to estimate density. It is a nonparametric statistics approach. It can used to estimate and fit density for a dataset, particularly for large size data. The Nigerian crime rate data reported to the Police was used to demonstrate the approach. This approach in the multivariate density estimation is based on the data. The quality of the estimates obtained of the MICI$_H$ approach showed some improvements over the existing methods. The MICI$_H$ approach has reduced mean squared error and relative faster rate of convergence compared to some other approaches. The approach of MICI$_H$ has reduced points of discontinuities in the graphical densities the datasets. They help to correct points of discontinuities and display adaptive density compared to the fixed $H$, MCKDE and MMCKDE approaches.

References


