Journal of Economics and Administrative Sciences
2024; 30 (139), pp. 126-148
P-ISSN 2518-5764
E-ISSN 2227-703X

Journal of Economics and Administrative Sciences (JEAS)

Available online at http://jeasiq.uobaghdad.edu.iq
DOI: https://doi.org/10.33095/tscrxh34

Probit and Improved Probit Transform-Based Kernel Estimator for Copula Density

Fatimah Hashim Falhi*
Department of Statistics,
College of Administration and Economics,
University of Basrah,
Basrah, Iraq.
fatma.hashem1101a@coadec.uobaghdad.edu.iq
fatima.falhi@uobasrah.edu.iq
http://orcid.org/0009-0008-1170-5635
***Corresponding author**

Munaf Youssif Hmood
Department of statistics,
College of Administration and Economics,
University of Baghdad.
Baghdad, Iraq
munaf.yousif@coadec.uobaghdad.edu.iq
http://orcid.org/0000-0002-1134-9078

Received:1/7/2023 **Accepted: 13/8/2023** **Published Online First: 29 /2/ 2024**

Abstract:

Copula modeling is widely used in modern statistics. The boundary bias problem is one of the problems faced when estimating by nonparametric methods, as kernel estimators are the most common in nonparametric estimation. In this paper, the copula density function was estimated using the probit transformation nonparametric method to eliminate of the boundary bias problem that suffers kernel estimators. Simulation was also employed for the for three nonparametric methods to estimate the copula density function and we proposed a new method that is better than the rest of the methods by five types of copulas with different sample sizes and different levels of correlation between the copula variables and the different parameters for the function. The results showed that the best method is to combine probit transformation and mirror reflection kernel estimator (PTMRKE) and followed by the (IPE) method when using all copula functions and for all sample sizes. If the correlation is strong (positive or negative). However, in the case of using weak and medium correlations, it turns out that the (IPE) method is the best, followed by the proposed method (PTMRKE), depending on (RMSE, LOGL, Akaike) criteria. The results also indicated weak mirror kernel reflection method when using the five copulas.

Paper type: Research paper.
Keywords: Copula function, probit transformation, Kernel copula function, Improved probit transformation, Mirror reflection, Boundary bias.

1. Introduction:

The nonparametric estimation technique is a common and flexible tool for analyzing data and modeling relationships between variables. The nonparametric estimation is different from the parametric estimation in that it does not take a fixed form or a specific form. Nevertheless, it is obtained according to the information derived from the data. All information regarding the phenomena under research is assumed to be regularly distributed in parametric models. Under tight assumptions and circumstances, we cannot use standard correlation measurements like Kendall's or Spearman's if the random variables are not normally distributed. Separating random variables' effects is extremely challenging, especially when evaluating the degree of positive and negative dependence. As a result, researchers use nonparametric approaches such as the kernel density function to detect dependencies, especially in multivariate distributions.

The problem in the modeling of multivariate functions is the presence of dependency between the observations of the variables of the examined phenomena, which can lead to various of issues, including boundary effects. In this situation, it is impossible to get the exact estimation for these functions. A suitable statistical tool must be used to characterize the dependence structure between the variables of the examined phenomenon, particularly when the effect extends over a long or medium period of time and the data distribution is unknown. Nonparametric approaches are employed to estimate the copula functions in this research.

Many studies have been published by researchers to help develop ideas for modeling dependency measures in many fields, especially the challenges encountered during the analysis, such as problems of association between study variables and problems of boundary effects.

Deheuvels (1979) developed the theory of nonparametric estimation of the copula function of a random variable based on the empirical copula and measured the sample dependency by employing of the empirical copula, and obtained a consistent empirical copula function.

Hmood (2005) clarified and reviewed some parametric, nonparametric, and semi-parametric methods and suggested methods for estimating the probability density function and choosing the appropriate method for estimating smoothing parameter and comparing the mentioned methods in determining the best estimator for the probability density function using the simulation method.

Dawod (2006) used the copula theory in modelling the survival function of the bivariate variable Weibull distribution and bivariate standard normal distribution cut off at zero point and using simulation experiments for comparison between the estimation of the survival function by using six different copula.

Genest and Favier (2007) presented a paper for inference copula models, based on the rank method. Working in detail on a small imaginary numeric example, illustrate the different steps for checking the dependence between two random variables and modeling it using copulas. It also introduces simple graphical tools and numerical techniques for selecting a suitable model, estimating its parameters, and checking its suitability. An application of the methodology to hydrological data is then presented.

Omelka et al. (2009) investigated kernel methods for obtaining smooth and flexible estimates of the bivariate correlation cumulative distribution function, and also discussed the selection of bandwidth parameters.

Chloob (2011) presented a proposal for a new copula by applying the Plackett copula through a mathematical modification that was made on that copula and comparing the Plackett copula with the proposed copula using simulations.

Geenens (2014) introduced the probit transformation of estimating the density of the kernel on the unit interval and he proposed a correct and simple method by combining the concept of transformation with estimating the local likelihood density, resulting in workable density estimations that are free of boundary issues in most cases.

Geenens et al. (2017) investigated the probit transformation of the nonparametric kernel estimation of the copula density. He proposed a kernel type copula density based on the idea of transforming the margin of copula density to normal distributions using the probit function and estimating the density in the transformed domain without boundary bias problems. Thus, obtaining an estimation of the copula density via the back-transformation, and it was then demonstrated that when this method is combined with methods of estimating the local polynomial density.

Hmood and Hamza (2019) presented a method for estimating the copula density using different kernel density methods, including the mirror reflection method, beta kernel method and kernel transformation method, and then comparing the three methods using simulation experiments, the results showed that The transformation kernel estimator is the best among the three methods, and it is proved that the copulas are highly explicitly for high dependency, especially of the Gaussian type.

Nagler (2021) presented a R package called Kdevine to estimate the density of the multivariate kernel with vine copulas.

Dawod (2022) studied reliability structural analysis methods with multidimensional correlation and when conducting a structural reliability analysis and calculating the probability of structural failure. The techniques that helped analyze structural reliability in light of the correlation problem, include the third-moment, fourth-moment, and D-Vine copula techniques. These techniques were based on the first-order reliability method in its basic techniques when transforming the studied random variables into independent standard normal random variables, and iterative algorithms were used to find the probability point of most failures.

These studies were confined to nonparametric kernel functions using a fixed-value smoothing coefficient or a symmetric diagonal matrix.

This research aims to estimate the copula density by nonparametric methods through probit transformation depending on the Kernel copula function to correct the boundary bias. Probit transformation is one of the methods used in boundary correction, and it is the most commonly used method, and because of what this method suffers from biases at boundary points, we used a smoothing coefficient in the form of a full positive matrix.

2. Materials and Methods:

2.1 Copula definition:

A copula is a function that illustrates modeling the dependency of random variables. Sklar created and initially utilized the copula in 1959.

This function has several advantages for modeling dependencies in multivariate data. First, consider the joint distribution's separation into the dependency structure (copula) and the basic marginal distributions.

And which can be viewed as a mathematical tool that is used to represent the relationship structure between two or more random variables. Many articles and studies have been written about the nonparametric estimation of copulas. Nonparametric methods are more flexible than standard parametric methods, as no assumptions are required.

According to Sklar theorem 1959, every joint cumulative distribution function F of continuous random quantities (X, Y) can be written as $F(x, y) = C(F_X(x), F_Y(y))$, for all $(x, y) \in R^2$, where F_X and F_Y are continuous marginal distributions and C: $[0, 1]^2 \rightarrow [0,1]$ is a unique corresponding to this joint distribution. Therefore, the copula is the joint cumulative distribution function with uniformly distributed marginal distributions on [0, 1] (Cherubini et al., (2004); Nelsen, (2006)).

Therefore, every multivariate CDFs with standard uniform marginal that show the dependence structure of random variables X and Y, and their marginal cumulative distribution functions are described by

P-ISSN 2518-5764
E-ISSN 2227-703X

$U = F_X(X)\ and\ V = F_Y(Y),$ (1)

where U and V are uniformly distributed variables and$(U, V) \in [0,1]$. The probability of two random variables, $X \leq x$ and $Y \leq y$, is described by the joint CDF $F_{XY}(X, Y) = P(X \leq x, Y \leq y)$.

$C(u, v) = \Pr(U \leq u, V \leq v),$ (2)

where $C(u, v)$ is called a copula and can be uniquely determined when u and v are continuous (Alsina et al., (2006)).
The following is the formula for a Gaussian copula: (Zeng et al., (2014))

$$C_\theta^{Ga}(u_1, u_2) = \frac{1}{2\pi\sqrt{1-\theta^2}} \int_{-\infty}^{\Phi^{-1}(u_1)} \int_{-\infty}^{\Phi^{-1}(u_2)} \exp[-\frac{s^2 - 2\theta st + t^2}{2(1-\theta^2)}] dtds$$

Φ Represents the standard normal distribution function, while Φ^{-1} represents the inverse of standard normal distribution function.
A Frank copula is given by (Chen and Guo, (2019))

$$C(u_1, u_2) = \frac{1}{\theta} log\left(1 + \frac{(e^{\theta u_1} - 1)(e^{\theta u_2} - 1)}{e^\theta - 1}\right), \theta \in (-\infty, +\infty)$$

Joe copula is provided by

$C_\alpha(u, v) = 1 - [(1-u)^\alpha + (1-v)^\alpha - (1-u)^\alpha(1-v)^\alpha]^{\frac{1}{\alpha}}$ as well as its density

$c_\alpha(u, v) = [w^\alpha + z^\alpha - wz^\alpha]^{\frac{1}{\alpha}-2} wz^{\alpha-1}[\alpha - 1 + w^\alpha + z^\alpha - wz^\alpha], \alpha \in [1, \infty).$

Where $w = 1 - u\ and\ z = 1 - v$. It is distinguished by upper tail dependency. moreover, $\lambda_U = 2 - 2^{\frac{1}{\alpha}}$. (André, (2019)).
Tawn copula is

$$C = exp\left\{(log(u_1) + log(u_2))A\left(\frac{log(u_2)}{log(u_1 u_2)}\right)\right\}, where$$

$$A(x) = (1-\alpha_1)x + (1-\alpha_2)(1-x) + \left((\alpha_1(1-x))^\theta + (\alpha_2 x)^\theta\right)^{\frac{1}{\theta}}$$

$and\ (\theta, \alpha_1, \alpha_2) \in (1, \infty) \times [0,1]^2$ $, for\ \alpha_1 = \alpha_2 = 1$, we recover the Gumbel copula.
$At\ any\ time\ \alpha_1 \neq \alpha_2$, it will be asymmetric in its components.

2.2 Kernel and probit estimation:

Numerous nonparametric methods exist for estimating the dependence structure between two random variables, such as polynomial approximation copulas and kernel smoothing copulas (Geenenes et al., (2017)).

2.2.1 Kernel density function estimation:

The d-dimensional multivariate kernel density estimator in its general form demonstrated by Hmood as bellow (Hmood et al., (2008); Gramacki (2018)).

$$\hat{f}(x, H) = \frac{1}{n|H|^{1/2}} \sum_{i=1}^{n} K(H^{-1/2}(x - X_i)) = \frac{1}{n} \sum_{i=1}^{n} K_H(x - X_i). \quad (3)$$

$$K_H(x) = |H|^{-1/2} K\left(H^{-1/2} x\right) \quad (4)$$

Where H is a positive and symmetric definite bandwidth matrix and K is a kernel function, and $|H| \to 0, n|H| \to \infty\ as\ n \to \infty$

There are several nonparametric techniques to estimate the dependence structure between two random variables, such as empirical. (Deheuvels, (1979)), polynomial approximation copula (Cherubini et al., (2004)) and kernel smoothing copulas (Charpentier et al., (2006); Cherubini et al., (2004)).

In the classical statistics texts, a kernel is a nonparametric method for estimating the probability density function (pdf) of a continuous random variable. Any probability density can be used for the kernel (Scott, (2009)).

In this study, we use kernel type copula estimators because this method is the most commonly used in the nonparametric estimation of copulas. Although its flexible (Geenenes,(2014)), But is not appropriate for the unit squared copula densities, essentially because it is heavily influenced by boundary bias issues for estimation function. In addition, most common copulas permit unbounded densities, and kernel methods are inconsistent in that case. Therefore, many researchers study and provide solutions to the boundary bias, including (Gijbels and Mielniczuk,(1990)) ;Charpentier et al., (2006) ; Geenens ,(2017)).

The standard kernel estimator for c, denoted by $\hat{c}^*$

$$\hat{c}^*(u,v) = \frac{1}{n|H_{UV}|^{1/2}} \sum_{i=1}^{n} K\left(H_{UV}^{-1/2}\begin{pmatrix} u - Ui \\ v - Vi \end{pmatrix}\right), \quad (5)$$

where $(u,v) \in [0,1]$and k: $R^2 \to R$. H_{UV} *is bandwidth matrix*

Using of kernel techniques to estimate an unknown bivariate copula density we will see that the boundedness of a copula density's support necessitates using of more advanced techniques than the one considered. U, V ~U[0, 1] are random variables with the joint distribution C and the corresponding density c: $[0, 1]^2 \to R$. We assume that the copula C has i.i.d variables $\{U_i = F_X(X_i), V_i = F_Y(Y_i)\ , i = 1, \dots, n\}$, and we aim to estimate the density c.**(**Genenes, (2014))

2.3 Probit Transformation Estimation Method (PTE):

Data transformations are commonplace and widely used to enhance the application and performance of classical estimating methods, this procedure deals with almost skewed data, heavy tails, or bounded support.

Several studies have investigated the transformation density estimation technique in the context of kernel density estimation, and they have presented several transformation families and transformation selection criteria. These studies created parametric families of transformations that approximate normality in a range of non-normal distributions. Although our essential goal of simple density estimation does not necessitate normality, Transformations can serve a variety of purposes in statistical analysis (Bean, (2017)).

To solve the problems that caused boundary bias by transforming the data to support its distribution on the full R^2.In other words, this method can be correct the boundaries in naturally, and this method is characterized by dealing with boundary copula densities (Charpentier et al., (2006)).

The difficulty in the copula density estimation of (U, V) is primarily due to the constrained nature of its support $= [0,1]^2$. Now define

$$S = \Phi^{-1}(U) \quad \text{and} \quad T = \Phi^{-1}(V). \quad (6)$$

Where Φ is the standard normal cumulative distribution function and Φ^{-1}its quantile function or the probit transformation. (Genenes, (2014) p5) Given that both U and V are uniform distributions [0,1], S and T have standard normal distributions, but this does not imply that the vector (S, T) is bivariate normal. If the joint CDF of (S,T) is the Gaussian, then F_{ST} is the

Gaussian copula because copulas are invariant for increasing transformations. (Nelsen, (2006), Theorem 2.4.3) (S, T) has unconstrained support R^2, and estimating its density f_{ST} cannot be affected by boundary problems. Furthermore, due to its normal margins, one can expect f_{ST} to be well-behaved and easy to estimate. Under mild assumptions, f_{ST} and its partial derivatives up to the second order are found to be bounded on R^2. In this case copula density is unbounded. If F_{ST} refer to copula C, and the variables (S,T) are standard normal distribution, then we can write Sklar's theorem as the equation below :

$$F_{ST}(s,t) = C\big(\Phi(s), \Phi(t)\big). \qquad (7)$$

When differentiating F_{ST} with respect to s and t, we get the joint density of (s,t),

$$f_{ST}(s,t) = c\big(\Phi(s), \Phi(t)\big)\varphi(s)\varphi(t), \qquad (8)$$

where φ is standard normal density. Inverting this equation yields.

$$c(u,v) = \frac{f_{ST}(\Phi^{-1}(u), \Phi^{-1}(v))}{\varphi(\Phi^{-1}(u))\varphi(\Phi^{-1}(v))}. \qquad (9)$$

For any $(u,v) \in [0,1]^2$, therefore, any estimator $\hat{f}_{ST}$ $on\ R^2$ automatically generates a Copula density estimate on the interior of I.

$$\hat{c}^{(\tau)}(u,v) = \frac{\hat{f}_{ST}(\Phi^{-1}(u), \Phi^{-1}(v))}{\varphi(\Phi^{-1}(u))\varphi(\Phi^{-1}(v))}, \qquad (10)$$

where the symbol (τ) refers to the transformation idea. When appropriate, $\hat{c}^{(\tau)}$ can alternatively be defined by continuity at the limits of I. This transformation-based estimator has a number of appealing qualities. Because $(\Phi^{-1}(u),\ \Phi^{-1}(v))$ is not defined for $(u,v) \notin I$ cannot allocate any probability outside I. Also, if f_{ST} is a true density function, in the sense that $f_{ST}\ (s,t) \geq 0\ for\ all\ (s,t)$and

$$\int\int_{R^2} \hat{f}_{ST}(s,t)dsdt = 1$$

Then, through transformation in variables $u = \Phi(s)$ and $v = \Phi(t)$,

$$\hat{c}^{(\tau)}(u,v) \geq 0\ \ forall\ \ u,v \in I \quad ; \int\int_I \hat{c}^{(\tau)}(u,v)dudv = 1$$

According to the bivariate kernel density estimator, which we shall denote by $\hat{f}_{ST}$. When apply to the copula:

$$c(u,v) = \frac{f_{ST}\big(\Phi^{-1}(u), \Phi^{-1}(v)\big)}{\varphi\big(\Phi^{-1}(u)\big)\varphi\big(\ \Phi^{-1}(v)\big)}, \qquad (11)$$

for all $u,v \in [0,1]^2$

The first basic idea is that we should use the standard kernel density estimator such as $\hat{f}_{ST}$. Specifically, we use the estimate as:

$$\hat{f}^{*}_{ST}(s,t) = \frac{1}{n|H_{ST}|^{1/2}} \sum_{i=1}^{n} K\left(H_{ST}^{-1/2}\begin{pmatrix} s - S_i \\ t - T_i \end{pmatrix}\right). \qquad (12)$$

Where K is a bivariate kernel function and $H_{ST}$is symmetric positive–definite matrix, and

$$\{S_i = \Phi^{-1}(U_i), T_i = \Phi^{-1}(V_i); \qquad i = 1, \dots, n\} \qquad (13)$$

is the transform domain sample. But (U_i, V_i) not available, and (S_i, T_i) as well. Instead, one must make use of

$$\left\{\left(\hat{S}_i = \Phi^{-1}(\widehat{U}_i), \hat{T}_i = \Phi^{-1}(\hat{V}_i)\right); i = 1, \dots, n\right\}. \quad (14)$$

That pseudo-transformed sample, as a result, the feasible form $\hat{f}^*_{ST}(s,t)$ is

$$\hat{f}_{ST}(s,t) = \frac{1}{n|H_{ST}|^{1/2}} \sum_{i=1}^{n} k\left(H_{ST}^{-1/2}\begin{pmatrix} s - \hat{S}_i \\ t - \hat{T}_i \end{pmatrix}\right). \quad (15)$$

Based on equation (11), this leads to a "probit transform kernel copula density estimator". (Genenes,G (2014) p5)

$$\hat{c}^{(\tau)}(u,v) = \frac{1}{n|H_{ST}|^{1/2}\varphi(\Phi^{-1}(u))\varphi(\Phi^{-1}(v))} \sum_{i=1}^{n} K\left(H_{ST}^{-1/2}\begin{pmatrix} \Phi^{-1}(u) - \Phi^{-1}(\widehat{U}_i) \\ \Phi^{-1}(v) - \Phi^{-1}(\hat{V}_i) \end{pmatrix}\right). \quad (16)$$

As a result, the asymptotic equation for the parameter of probit transformation is also obtained. The bias and variance of this method for copula density estimator are in the following form, respectively.

$$Bias[\hat{c}^{\tau}(u,v)] = \frac{1}{2} m_2(K)\left\{h_{11}\left[c_{uu}(u,v)\varphi^2\left(\Phi^{-1}(u)\right) - 3c_u(u,v)\varphi\left(\Phi^{-1}(u)\right)\Phi^{-1}(u) + c(u,v)\left(\left(\Phi^{-1}(u)\right)^2 - 1\right)\right] + h_{22}\left[c_{vv}(u,v)\Phi^{-1}(u)\Phi^{-1}(v) - 3c_v(u,v)\varphi\left(\Phi^{-1}(v)\right)\Phi^{-1}(v) + c(u,v)\left(\left(\Phi^{-1}(v)\right)^2 - 1\right)\right] + 2h_{12}\left[c(u,v)\Phi^{-1}(u)\Phi^{-1}(v) - c_u(u,v)\varphi\left(\Phi^{-1}(u)\right)\Phi^{-1}(v) - c_v(u,v)\Phi^{-1}(u)\varphi\left(\Phi^{-1}(v)\right)\right]\right\} + o\{tr(H)\}. \quad (17)$$

Where $m_2(K) = \int z^2\, K(z)dz$

The variance is

$$var\left(\hat{f}(s,t)\right) = n^{-1}|H|^{-1/2}R(K)f(s,t) + o\left(n^{-1}|H|^{-1/2}\right). \quad (18)$$

Where $R(K) = \int K(z)^2\, dz$

Then the variance of probit transformation copula density as below

$$var\left(\hat{c}^{\tau}(u,v)\right) = \frac{R(K)}{n|H|^{1/2}} \times \frac{c(u,v)}{\varphi(\Phi^{-1}(u))\varphi(\Phi^{-1}(v))} + o((n|H|)^{-1}). \quad (19)$$

When we use standard normal distribution of kernel density and normal distribution for density function then,

$m_2(K) = 1$ and $R(k) = (4\pi)^{-d/2}$. Where d represents a number of variables

Observe that

$$H_{AMISE} = \left[\frac{4}{(d+2)n}\right]^{\frac{2}{d+4}} \hat{\Sigma}. \quad (20)$$

2.4 Improved probit transformation method (IPT):

An extension of transformation method is proposed by (Geenens, (2014)) fitting local polynomial to the log-density for the sample transformation and by quadratic polynomials.

The purpose and the advantages of estimating f_{ST} by the local likelihood methods as an alternative to standard kernel density estimation is related to the boundary behavior of the estimator of c on I and the tail behavior of the estimator of f_{ST} on R^2. But, standard kernel estimators are well-known to work unsuccessfully in the tails of densities, with repeated occurrences of 'spurious bumps'. These fluctuations are greatly magnified by back transformation (11), the so- yielded estimator of c illustrations a very irregular behavior at the boundaries.

The local likelihood technique (Loader, (1996)) assumes that the log-density $log\, f(s,t)$ of the random vector $(s,t) = (\Phi^{-1}(U), \Phi^{-1}(V))$ may be approximated locally by a polynomial $P_a(s,t)$ of order p. The coefficient vector of the polynomial is denoted as $a(s,t) \in R\ (p+1)(p+2)/2$, where $(p+1)(p+2)/2$ is simply the number of terms (including a constant) of a two- dimensional polynomial of order P. Then we can write local likelihood estimator as follows in this context of estimating f_{ST} from the pseudo-sample $(\hat{S}_i, \hat{T}_i), i = 1, \dots, n.$ $logf_{ST}$ is assumed to be well approximated by a polynomial of order p about $(s,t) \in R^2$. Only local log-linear (p = 1) and local log-quadratic (p = 2) estimators are studied classically. In particular, in the first order (p = 1), it is assumed that, given $(\tilde{s}, \tilde{t})$ 'converge' to (s, t),

Local log linear (p=1) it is follow:

$$logf_{ST}(\tilde{s},\tilde{t}) \cong a_{10}(s,t) + a_{11}(s,t)(\tilde{s}-s) + a_{12}(s,t)(\tilde{t}-t) = P_{a_i}(\tilde{s}-s, \tilde{t}-t). \quad (21)$$

In the second order (p = 2),

$$logf_{ST}(\tilde{s},\tilde{t}) \cong a_{20}(s,t) + a_{21}(\tilde{s}-s) + a_{22}(\tilde{t}-t) + a_{23}(s,t)(\tilde{s}-s)^2 + a_{24}(s,t)(\tilde{t}-t)^2 + a_{25}(\tilde{s}-s)(\tilde{t}-t) = P_{a_2}(\tilde{s}-s, \tilde{t}-t). \quad (22)$$

The vectors

$$a_1(s,t) = \left(a_{10}(s,t), a_{11}(s,t), a_{12}(s,t)\right)\ and\ \ a_2(s,t) = (a_{20}(s,t), \dots, a_{25}(s,t). \quad (23)$$

By solving a weighted maximum likelihood problem. For either p=1or P=2

$$\tilde{a}_p(s,t) = argmax\left\{\sum_{i=1}^{n} K\left(H_{ST}^{-1/2}\begin{pmatrix} s-\hat{s} \\ t-\hat{t}\end{pmatrix}\right) P_{a_p}(\hat{s}-s)(\hat{t}-t) - n\int\int K\left(H_{ST}^{-1/2}\begin{pmatrix} s-\tilde{s} \\ t-\tilde{t}\end{pmatrix}\right)\exp(P_{a_p}(\tilde{s}-s,\tilde{t}-t)d\tilde{s}d\tilde{t}\right\}. \quad (24)$$

Where K is a bivariate kernel function and H_{ST} is a symmetric positive-definite bandwidth matrix, as previously stated.

The improved probit transformation estimation for kernel copula density. In the case of the local log-linear (p = 1)

$$\tilde{c}^{(\tau,1)}(u,v) = \frac{\exp(\tilde{a}_{1,0}\left(\Phi^{-1}(u), \Phi^{-1}(v)\right))}{\varphi\left(\Phi^{-1}(u)\right)\varphi\left(\Phi^{-1}(v)\right)}. \quad (25)$$

And in the case of the local log-quadratic (p=2) approximation

$$\tilde{c}^{(\tau,2)}(u,v) = \frac{\exp(\tilde{a}_{2,0}(\Phi^{-1}(u),\Phi^{-1}(v)))}{\varphi(\Phi^{-1}(u))\varphi(\Phi^{-1}(v))}. \quad (26)$$

We get for all $(s,t) \in R^2$ where $f_{ST}(s,t)$ is positive and continuous second-order partial derivatives are admissible approximation estimator $\tilde{f}_{ST}^{(1)}$ to calculate the joint density f_{ST}. (Loarder, (1996))
Define the optimum local log-linear probit-transformation kernel copula density estimator.

$$\tilde{c}^{*(\tau,1)}(u,v) = \frac{\tilde{f}_{ST}^{*(\tau,1)}(s,t)}{\varphi(\Phi^{-1}(u))\varphi(\Phi^{-1}(v))}. \quad (27)$$

Using (9), (11) and (19) in (27), one obtains.

$$\sqrt{nh^2}\left(\tilde{c}^{*(\tau,1)}(u,v) - c(u,v) - h^2 b^{(1)}(u,v)\right) \xrightarrow{L} N\left(0, {\sigma^{(1)}}^2(u,v)\right). \quad (28)$$

Then the bias local linear probit transformation equals the equation

$$IB(c^{(\tau,1)}(u,v)) = h_{11}\left\{\left[\frac{\partial^2 c(u,v)}{\partial u^2}\varphi^2(\Phi^{-1}(u)) - \frac{\partial c(u,v)}{\partial u}\Phi^{-1}(u)\left(\varphi(\Phi^{-1}(u))\right) - \frac{1}{c(u,v)}\left[\frac{\partial c(u,v)}{\partial u}\right]^2 \varphi^2(\Phi^{-1}(u))\right] - c(u,v)\right\} + 2h_{12}\left\{\left[\frac{\partial^2 c(u,v)}{\partial u \partial v}\varphi(\Phi^{-1}(u))\varphi(\Phi^{-1}(v)) - \frac{1}{c(u,v)}\left(\frac{\partial c(u,v)}{\partial u}\varphi(\Phi^{-1}(u))\right)\left(\frac{\partial c(u,v)}{\partial v}\varphi(\Phi^{-1}(v))\right)\right]\right\} + h_{22}\left\{\left[\frac{\partial^2 c(u,v)}{\partial u \partial v}\varphi(\Phi^{-1}(u))\varphi(\Phi^{-1}(v)) - \frac{1}{c(u,v)}\left(\frac{\partial c(u,v)}{\partial u}\varphi(\Phi^{-1}(u))\right)\left(\frac{\partial c(u,v)}{\partial v}\varphi(\Phi^{-1}(v))\right)\right]\right\}. \quad (29)$$

And the variance

$$Ivar(c^{(\tau,1)}(u,v)) = \frac{c(u,v)}{4\pi|H|^{1/2}\varphi(\Phi^{-1}(u))\varphi(\Phi^{-1}(v))}. \quad (30)$$

Local log-quadratic probit-transformation kernel copula density estimator $\tilde{c}^{(\tau,2)}$ for all $(u,v) \in (0,1)^2$ is such that

$$\sqrt{nh^2}\left(\tilde{c}^{*(\tau,2)}(u,v) - c(u,v) - h^4 b^{(2)}(u,v)\right) \xrightarrow{L} N\left(0, {\sigma^{(2)}}^2(u,v)\right). \quad (31)$$

Where

$${\sigma^{(2)}}^2(u,v) = \frac{5}{2}\,\frac{c(u,v)}{4\pi|H|^{1/2}\varphi(\Phi^{-1}(u))\varphi(\Phi^{-1}(v))}. \quad (32)$$

And $b^{(2)}(u,v)$ is a similar equation to (probit transformation), except it involves partial derivatives of c up to the fourth order.

2.5 Mirror Reflection Kernel Estimation Method (MRKE):

Kernel estimation for copula is famous for suffering from boundary bias. One technique of removing this difficulty is by reflecting all data points with regard to each corner and edge of the unit square (Charpentier et al., 2006; Nagler, 2014). This idea was presented by (Gijbels and Mielniczuk, 199). And the method is known as mirror reflection. This procedure aims to add some "missing mass" to the sample by reflecting it with regard to the boundaries. They concentrate on the scenario where the variables are positive and have support as $[0,\infty)$.

The mirror reflection kernel takes the form (Gijbels and Mielniczuk,(1990)).

$$\hat{c}_n^{(MR)} = \frac{1}{n}\sum_{i=1}^{n}\sum_{l=1}^{9}\left[K\left(\frac{u-\hat{U}_{il}}{h_n}\right)-K\left(\frac{-\hat{U}_{il}}{h_n}\right)\right]\left[K\left(\frac{v-\hat{V}_{il}}{h_n}\right)-K\left(\frac{-\hat{V}_{il}}{h_n}\right)\right]. \quad (33)$$

With $\{(\hat{U}_{il}),(\hat{V}_{il})\} = \{.(\pm\hat{U}_i,\pm\hat{V}_i),(\pm\hat{U}_i,2-\hat{V}_i),(2-\hat{U}_i,\pm\hat{V}_i),(2-\hat{U}_i,2-\hat{V}_i)\}$ (Charpentier et al., 2006).

An estimated formula for the reflection density function of the copula mirror can be written as

$$\hat{c}_h^{(MR)}(u,v) = \frac{1}{nh^2}\sum_{i=1}^{n}\Big\{K\left(\frac{u-\hat{U}_i}{h}\right)K\left(\frac{v-\hat{V}_i}{h}\right)+K\left(\frac{u+\hat{U}_i}{h}\right)K\left(\frac{v-\hat{V}_i}{h}\right)+K\left(\frac{u-\hat{U}_i}{h}\right)K\left(\frac{v+\hat{V}_i}{h}\right)+ K\left(\frac{u+\hat{U}_i}{h}\right)K\left(\frac{v+\hat{V}_i}{h}\right)+K\left(\frac{u-\hat{U}_i}{h}\right)K\left(\frac{v-2+\hat{V}_i}{h}\right)+K\left(\frac{u+\hat{U}_i}{h}\right)K\left(\frac{v-2+\hat{V}_i}{h}\right)+K\left(\frac{u-2+\hat{U}_i}{h}\right)K\left(\frac{v-\hat{V}_i}{h}\right)+ K\left(\frac{u-2+\hat{U}_i}{h}\right)K\left(\frac{v+\hat{V}_i}{h}\right)+ K\left(\frac{u-2+\hat{U}_i}{h}\right)K\left(\frac{v-2+\hat{V}_i}{h}\right)\Big\}. \quad (34)$$

When we use full bandwidth matrix H the mirror reflection copula estimator as

$$\hat{c}^{(MR)}(u,v) = \frac{1}{n|H|^{1/2}}\sum_{i=1}^{n}\sum_{k=1}^{9}K(H^{-1/2}\begin{pmatrix}u-\hat{U}_{ik}\\ v-\hat{V}_{ik}\end{pmatrix}). \quad (35)$$

Then the bias is the following formula

$$Bias\left(\hat{c}^{(MR)}(u,v)\right) = \frac{1}{2}m_2(K)tr\{H\mathcal{H}_c(u,v)\}+o\{tr(H)\}$$

$$where\ H = \begin{bmatrix}h_{11} & h_{12}\\ h_{12} & h_{22}\end{bmatrix}\ and\ \mathcal{H}_c(u,v) = \begin{bmatrix}c_{uu} & c_{uv}\\ c_{uv} & c_{vv}\end{bmatrix}.$$

Then $ABias\left(\hat{c}^{(MR)}(u,v)\right) = \frac{1}{2}m_2(K)\{h_{11}c_{uu}+2h_{12}c_{uv}+h_{22}c_{vv}\}$, (36)

c_{uu} *and* c_{vv} are the second derivatives for u and v respectively c_{uv} is the mixed second derivative

The variance formula is as shown below:

$$var\left(\hat{c}^{(MR)}(u,v)\right) = \frac{1}{n|H|^{1/2}}R(K)c(u,v)+o\left(\frac{1}{n|H|^{1/2}}\right). \quad (37)$$

2.6 Bandwidth selection:

The problem of selecting the bandwidth parameter is a crucial problem that occurs often in the context of KDE. The precision of KDE depends on the chosen bandwidth value. In the univariate case, the bandwidth is a scalar controlling the smoothing quantity. In the multivariate case, the bandwidth is a matrix that controls both the quantity and the smoothing shape. This matrix can be defined on various levels of complexity. (Gramacki, 2018)

The bandwidth affects the balance between two concerns in nonparametric estimation: bias and variance. Furthermore, the mean squared error (MSE), which is the sum of squared bias and variance, performs composite metric. As a result, optimality in the sense of MSE is not significantly influenced by the kernel selection but is influenced by the bandwidth selection (Bowman, 1997).There are several techniques for calculating the bandwidth h. The plug-in approach and cross validation are two of the most often used. We utilize Silverman's rule of thumb bandwidth h for the plug-in approach in all methods and every sample size. We used plug–in method for the selection bandwidth matrix for all methods.

2.7 Performance Criteria:

The comparison between the estimation methods is carried out according to the Root Mean Squares Error (RMSE) and is done by calculating the mean squares error of the copula function estimated for each iteration according to the following formula:-

$MSE(\hat{c}_H, c) = E(\hat{c}_H(u,v) - c(u,v))^2$

$RMSE(\hat{c}_H, c) = \sqrt{EE(\hat{c}_H(u,v) - c(u,v))^2}.$

And the Akaike criterion (AIC) is:

$$AIC_n^{(.)} := -2\sum_{i=1}^{n} \ln\left(c_{\theta_n}^{(.)}\left(u_1^{(i)}, \dots, u_1^{(i)}\right)\right) + 2p,$$

where p is the number of family parameters and θn is a parameter estimate.The logarithm of maximum likelihood possibility (LOG L).

$L(\theta;\ u_1, \dots, u_n) = \prod_{i=1}^{n} c_\theta(u_i) \quad and\ l(\theta; u_1, \dots, u_n) = \sum_{i=1}^{n} l(\theta; u_i),$

respectively, where:

$l(\theta; u_i) = lnc_\theta(u_i) = ln[(-1)^d {\psi_\theta}^d[t_\theta(u_i)] + \sum_{j=1}^{d} \ln\left[-\left(\psi_\theta^{-1}\right)'(u_i)\right].$

The best method is the one that minimize root mean square error and minimize information criterion, both criteria select the model that gives the highest likelihood.

3. Discussion of results:

Simulation experiments were carried out using five sample sizes (n = 32, 64, 128, 256, and 512) with a frequency of 1,000 for each experiment, as follows:

1- The variables u and v are distributed uniformly.

2- Finding the probit transformation of the observations of the variables that were generated in step 1.

3-Five copulas of Gaussian, Frank, Tawn, RotationTawn and Joe were used depending on the different values of each copula parameter.

Tables 1 to 15 represent the estimated root mean squares error of the copula density functions for nonparametric estimation methods and Akaike criteria and logarithm maximum likelihood criteria (LogL) at a correlation level tau = 0.7,0.5,0.3 respectively with 1000 repetitions for each experiment that were used to determine the performance of the best estimation method it was found that the best estimation method for the copula density function in the case of strong negative and positive correlations and for all sample sizes and for five copulas(Gaussian, Frank, Tawn, RTawn, and Joe) it is the proposed method (PTMRKE) followed by the improved probit transformation method(IPE), but in the case of medium and weak correlations, the best estimation method is the improved probit transformation method(IPE), followed by the proposed method (PTMRKE) when using the five copulas and for all sample sizes. The third method was probit transformation for all sample sizes and for all five copulas. The fourth and last place was the mirror reflection method (MRKE) for all sample sizes and copula functions.

Table 1: Root-mean square error, (AIC) criterion and logarithm likelihood criteria for Gaussian copula when $\tau = 0.7$.

Gaussian / Sample size	Method	RMSE	AIC	LOGL
32	PTE	0.29933	-38.599	22.71667
	IPE	0.21762	-48.7999	25.68059
	MRKE	0.36543	-32.2141	17.57557
	PTMRKE	0.16616	-70.6815	36.17277
64	PTE	0.23146	-97.1761	49.99241
	IPE	0.19774	-115.534	58.9331
	MRKE	0.32671	-71.5952	40.34811
	PTMRKE	0.16125	-127.243	64.63423
128	PTE	0.21907	-215.913	109.6572
	IPE	0.18834	-246.026	124.3943
	MRKE	0.2748	-199.168	106.4232
	PTMRKE	0.14509	-283.616	142.8579
256	PTE	0.22168	-374.616	194.7401
	IPE	0.16444	-427.389	215.3725
	MRKE	0.25644	-330.682	167.3907
	PTMRKE	0.10983	-466.362	234.6158
512	PTE	0.18511	-771.349	387.8594
	IPE	0.1463	-829.842	416.7956
	MRKE	0.21098	-696.762	356.7875
	PTMRKE	0.10907	-854.544	428.9544

Table 2: Root-mean square error, (AIC) criterion and logarithm likelihood criteria for Gaussian copula when $\tau = 0.5$.

Gaussian / Sample size	Method	RMSE	AIC	LOGL
32	PTE	0.64513	-14.0259	9.10046
	IPE	0.4437	-27.6558	15.38779
	MRKE	0.68337	-12.1829	8.23528
	PTMRKE	0.46592	-26.3567	14.71949
64	PTE	0.51196	-50.1634	27.14061
	IPE	0.42137	-81.6028	42.28345
	MRKE	0.5836	-36.0855	20.84349
	PTMRKE	0.43883	-67.5245	35.44688
128	PTE	0.49618	-102.509	53.73536
	IPE	0.36003	-130.761	67.39552
	MRKE	0.57604	-74.4768	41.22462
	PTMRKE	0.36488	-129.049	66.47255
256	PTE	0.44895	-180.244	95.50428
	IPE	0.34212	-271.015	137.6494
	MRKE	0.56902	-137.868	71.95402
	PTMRKE	0.35817	-267.205	135.8615
512	PTE	0.42957	-342.784	174.5219
	IPE	0.27942	-416.055	210.6041
	MRKE	0.49456	-282.411	146.4054
	PTMRKE	0.34692	-391.041	198.1826

P-ISSN 2518-5764
E-ISSN 2227-703X

Table3: Root-mean square error, (AIC) criterion and logarithm likelihood criteria for Gaussian copula when $\tau = 0.3$.

Gaussian Sample size	Method	RMSE	AIC	LOGL
32	PTE	0.90599	-10.0097	7.17241
	IPE	0.66802	-19.3517	11.51404
	MRKE	0.95078	-3.40293	3.41731
	PTMRKE	0.80489	-10.3747	7.26159
64	PTE	0.80242	-25.8767	15.42052
	IPE	0.56084	-47.4618	25.66639
	MRKE	0.85514	-10.0902	7.04017
	PTMRKE	0.79767	-26.5584	15.67801
128	PTE	0.73139	-27.7824	17.43187
	IPE	0.54622	-113.279	58.593
	MRKE	0.85414	-6.46011	4.65237
	PTMRKE	0.72583	-12.6367	10.3235
256	PTE	0.71448	-86.0686	46.49385
	IPE	0.47636	-148.05	76.6885
	MRKE	0.81988	-50.8407	28.00765
	PTMRKE	0.7096	-87.8233	47.25986
512	PTE	0.66743	-112.706	60.65348
	IPE	0.41272	-206.561	106.5277
	MRKE	0.89912	-80.1228	42.69741
	PTMRKE	0.66036	-115.742	62.20851

Table 4: Root-mean square error, (AIC) criterion and logarithm likelihood criteria for Frank copula when $\tau = 0.7$.

Frank Sample size	Method	RMSE	AIC	LOGL
32	PTE	0.15321	-81.6971	41.67683
	IPE	0.11183	-92.0527	46.69053
	MRKE	0.33162	-74.2142	37.9642
	PTMRKE	0.09322	-106.557	53.76981
64	PTE	0.15021	-95.0123	48.9454
	IPE	0.11156	-113.198	57.75536
	MRKE	0.27301	-70.6406	39.20701
	PTMRKE	0.07785	-119.95	61.03871
128	PTE	0.14841	-280.421	141.4846
	IPE	0.10953	-303.719	152.9268
	MRKE	0.26035	-218.77	110.8423
	PTMRKE	0.07399	-333.29	167.5095
256	PTE	0.14168	-485.297	244.3296
	IPE	0.09936	-535.622	269.214
	MRKE	0.23254	-344.508	174.2283
	PTMRKE	0.05632	-561.451	281.9717
512	PTE	0.12341	-895.484	449.935
	IPE	0.09835	-976.766	490.1482
	MRKE	0.2266	-746.853	383.1753
	PTMRKE	0.0358	-1182.2	592.7062

Table 5: Root-mean square error, (AIC) criterion and logarithm likelihood criteria for Frank copula when $\tau = 0.5$.

Frank / Sample size	Method	RMSE	AIC	LOGL
32	PTE	0.42059	-16.5343	10.4112
	IPE	0.37463	-18.7305	11.3339
	MRKE	0.8569	-4.94191	4.45647
	PTMRKE	0.31399	-28.4268	15.93552
64	PTE	0.41134	-59.704	31.66972
	IPE	0.3631	-64.8368	34.18017
	MRKE	0.77787	-31.5412	18.83118
	PTMRKE	0.29965	-77.1296	40.05924
128	PTE	0.39342	-148.273	76.18602
	IPE	0.34798	-179.281	91.29146
	MRKE	0.61089	-109.55	59.2531
	PTMRKE	0.24644	-188.787	95.93041
256	PTE	0.38824	-221.718	113.3838
	IPE	0.33548	-251.344	127.829
	MRKE	0.59649	-156.59	82.5269
	PTMRKE	0.2288	-270.235	137.1816
512	PTE	0.38815	-484.543	245.4771
	IPE	0.32608	-573.754	289.3423
	MRKE	0.55692	-368.257	189.9977
	PTMRKE	0.42492	-559.206	282.0944

Table 6: Root-mean square error, (AIC) criterion and logarithm likelihood criteria for Frank copula when $\tau = 0.3$.

Frank / Sample size	Method	RMSE	AIC	LOGL
32	PTE	0.90599	-10.0097	7.17241
	IPE	0.66802	-19.3517	11.51404
	MRKE	0.95078	-3.40293	3.41731
	PTMRKE	0.80489	-10.3747	7.26159
64	PTE	0.80242	-25.8767	15.42052
	IPE	0.56084	-47.4618	25.66639
	MRKE	0.85514	-10.0902	7.04017
	PTMRKE	0.79767	-26.5584	15.67801
128	PTE	0.73139	-27.7824	17.43187
	IPE	0.54622	-113.279	58.593
	MRKE	0.85414	-6.46011	4.65237
	PTMRKE	0.72583	-12.6367	10.3235
256	PTE	0.71448	-86.0686	46.49385
	IPE	0.47636	-148.05	76.6885
	MRKE	0.81988	-50.8407	28.00765
	PTMRKE	0.7096	-87.8233	47.25986
512	PTE	0.66743	-112.706	60.65348
	IPE	0.41272	-206.561	106.5277
	MRKE	0.89912	-80.1228	42.69741
	PTMRKE	0.66036	-115.742	62.20851

P-ISSN 2518-5764
E-ISSN 2227-703X

Table 7: Root-mean square error, (AIC)criterion and logarithm likelihood criteria for Tawn copula when $\tau = 0.7$.

Tawn Sample size	Method	RMSE	AIC	LOGL
32	PTE	0.15321	-81.6971	41.67683
	IPE	0.11183	-92.0527	46.69053
	MRKE	0.33162	-74.2142	37.9642
	PTMRKE	0.09322	-106.557	53.76981
64	PTE	0.15021	-95.0123	48.9454
	IPE	0.11156	-113.198	57.75536
	MRKE	0.27301	-70.6406	39.20701
	PTMRKE	0.07785	-119.95	61.03871
128	PTE	0.14841	-280.421	141.4846
	IPE	0.10953	-303.719	152.9268
	MRKE	0.26035	-218.77	110.8423
	PTMRKE	0.07399	-333.29	167.5095
256	PTE	0.14168	-485.297	244.3296
	IPE	0.09936	-535.622	269.214
	MRKE	0.23254	-344.508	174.2283
	PTMRKE	0.05632	-561.451	281.9717
512	PTE	0.12341	-895.484	449.935
	IPE	0.09835	-976.766	490.1482
	MRKE	0.2266	-746.853	383.1753
	PTMRKE	0.0358	-1182.2	592.7062

Table 8: Root-mean square error, (AIC) criterion and logarithm likelihood criteria for Tawn copula when $\tau = 0.5$.

Tawn Sample size	Method	RMSE	AIC	LOGL
32	PTE	0.42059	-16.5343	10.4112
	IPE	0.37463	-18.7305	11.3339
	MRKE	0.8569	-4.94191	4.45647
	PTMRKE	0.31399	-28.4268	15.93552
64	PTE	0.41134	-59.704	31.66972
	IPE	0.3631	-64.8368	34.18017
	MRKE	0.77787	-31.5412	18.83118
	PTMRKE	0.29965	-77.1296	40.05924
128	PTE	0.39342	-148.273	76.18602
	IPE	0.34798	-179.281	91.29146
	MRKE	0.61089	-109.55	59.2531
	PTMRKE	0.24644	-188.787	95.93041
256	PTE	0.38824	-221.718	113.3838
	IPE	0.33548	-251.344	127.829
	MRKE	0.59649	-156.59	82.5269
	PTMRKE	0.2288	-270.235	137.1816
512	PTE	0.38815	-484.543	245.4771
	IPE	0.32608	-573.754	289.3423
	MRKE	0.55692	-368.257	189.9977
	PTMRKE	0.42492	-559.206	282.0944

Table 9: Root-mean square error, (AIC) criterion and logarithm likelihood criteria for Tawn copula when $\tau = 0.3$.

Tawn Sample size	Method	RMSE	AIC	LOGL
32	PTE	0.7151	-9.30299	6.94363
	IPE	0.49316	-59.4376	30.6466
	MRKE	0.98378	-0.52327	1.79105
	PTMRKE	0.82096	-3.07841	3.91647
64	PTE	0.71314	-21.1549	13.43301
	IPE	0.41848	-42.2807	23.35492
	MRKE	0.9348	-4.31536	3.6875
	PTMRKE	0.77409	-12.6261	9.41658
128	PTE	0.6932	-56.5897	31.11079
	IPE	0.41745	-90.4873	47.44672
	MRKE	0.90828	-12.8533	8.34424
	PTMRKE	0.74773	-46.9046	26.40814
256	PTE	0.69094	-104.666	55.57158
	IPE	0.56379	-146.708	76.11861
	MRKE	0.90446	-65.5524	35.51815
	PTMRKE	0.72871	-99.2848	53.04566
512	PTE	0.67853	-171.691	90.04039
	IPE	0.51652	-447.28	226.4952
	MRKE	0.86144	-76.3629	40.78896
	PTMRKE	0.71853	-162.035	85.20885

Table 10: Root-mean square error, (AIC) criterion and logarithm likelihood criteria for Rotation Tawn copula when $\tau = 0.7$.

RTawn Sample size	Method	RMSE	AIC	LOGL
32	PTE	0.23397	-57.2149	33.03649
	IPE	0.22434	-70.0372	35.94978
	MRKE	0.32936	-47.8263	25.121
	PTMRKE	0.18981	-94.4901	47.82346
64	PTE	0.16509	-115.282	58.98329
	IPE	0.16245	-131.871	67.01851
	MRKE	0.16823	-71.0787	37.31199
	PTMRKE	0.16126	-149.893	75.81619
128	PTE	0.1342	-308.368	155.2586
	IPE	0.12739	-316.724	159.3168
	MRKE	0.15286	-228.148	120.5535
	PTMRKE	0.12124	-323.996	162.9021
256	PTE	0.10253	-540.165	271.6675
	IPE	0.10036	-583.067	292.8519
	MRKE	0.11207	-467.14	242.3051
	PTMRKE	0.09613	-617.007	309.5827
512	PTE	0.08421	-930.513	467.163
	IPE	0.07956	-968.129	485.7099
	MRKE	0.08989	-731.277	375.7
	PTMRKE	0.05219	-1177.07	589.6745

Table 11: Root-mean square error, (AIC) criterion and logarithm likelihood criteria for Rotation Tawn copula when $\tau = 0.5$.

RTawn Sample size	Method	RMSE	AIC	LOGL
32	PTE	0.60703	-26.265	14.88072
	IPE	0.53565	-41.3756	21.96108
	MRKE	0.61915	-16.5941	11.07809
	PTMRKE	0.59537	-28.3714	15.86527
64	PTE	0.48253	-46.6379	25.57054
	IPE	0.43968	-64.2414	33.93003
	MRKE	0.48314	-30.6823	18.04697
	PTMRKE	0.46315	-58.5933	31.15594
128	PTE	0.41205	-98.5228	51.86625
	IPE	0.39201	-130.591	67.33317
	MRKE	0.42299	-63.8942	35.30246
	PTMRKE	0.40031	-120.508	62.37158
256	PTE	0.38617	-186.226	96.1536
	IPE	0.36649	-243.575	124.1609
	MRKE	0.39005	-133.729	70.95541
	PTMRKE	0.36879	-220.266	112.6513
512	PTE	0.29948	-420.718	213.685
	IPE	0.21446	-510.732	257.9583
	MRKE	0.32333	-319.628	165.4794
	PTMRKE	0.27445	-479.478	242.4286

Table 12: Root-mean square error, (AIC) criterion and logarithm likelihood criteria for Rotation Tawn copula when $\tau = 0.3$.

RTawn Sample size	Method	RMSE	AIC	LOGL
32	PTE	0.85727	-20.5896	12.18931
	IPE	0.81826	-29.9002	16.51437
	MRKE	0.97197	-1.81957	2.45805
	PTMRKE	0.82609	-27.5715	15.40793
64	PTE	0.71005	-19.5586	12.57756
	IPE	0.67178	-39.2835	21.88297
	MRKE	0.74568	-9.39085	7.02143
	PTMRKE	0.68581	-36.3116	20.36863
128	PTE	0.62866	-59.4131	32.42453
	IPE	0.59029	-67.4736	36.34011
	MRKE	0.66807	-19.7116	12.08966
	PTMRKE	0.63266	-58.6007	32.09161
256	PTE	0.54129	-96.5501	51.66227
	IPE	0.53813	-154.619	79.97331
	MRKE	0.55765	-35.7064	20.04867
	PTMRKE	0.54663	-87.658	47.34806
512	PTE	0.42248	-206.519	106.9281
	IPE	0.39283	-299.016	152.4928
	MRKE	0.47904	-102.692	54.40671
	PTMRKE	0.42702	-203.686	105.7278

P-ISSN 2518-5764
E-ISSN 2227-703X

Table 13: Root-mean square error, (AIC) criterion and logarithm likelihood criteria for Joe copula when $\tau = 0.7$.

JOE Sample size	Method	RMSE	AIC	LOGL
32	PTE	0.17041	-86.6256	43.92094
	IPE	0.12401	-86.8545	44.07575
	MRKE	0.22733	-80.0315	40.74686
	PTMRKE	0.082	-106.078	53.52986
64	PTE	0.15969	-117.411	59.98481
	IPE	0.123	-136.165	69.12226
	MRKE	0.21424	-98.6034	54.56637
	PTMRKE	0.07385	-166.765	84.12771
128	PTE	0.15494	-214.156	108.7846
	IPE	0.12205	-267.191	134.7292
	MRKE	0.20246	-188.59	99.29476
	PTMRKE	0.06673	-268.236	135.313
256	PTE	0.14883	-453.003	228.3562
	IPE	0.11913	-498.865	250.9384
	MRKE	0.18766	-408.883	211.4487
	PTMRKE	0.06626	-538.458	270.4771
512	PTE	0.14246	-872.756	438.4459
	IPE	0.11839	-954.745	479.0451
	MRKE	0.18184	-758.458	387.4757
	PTMRKE	0.03634	-975.336	489.1793

Table 14: Root-mean square error, (AIC) criterion and logarithm likelihood criteria for Joe copula when $\tau = 0.5$.

JOE Sample size	Method	RMSE	AIC	LOGL
32	PTE	0.42297	-20.1695	11.94385
	IPE	0.32579	-30.0756	16.57662
	MRKE	0.75801	-8.86603	6.2354
	PTMRKE	0.37517	-23.6385	13.54153
64	PTE	0.41975	-55.5584	29.87232
	IPE	0.30208	-71.777	37.61576
	MRKE	0.58183	-34.6611	20.29781
	PTMRKE	0.35761	-71.7391	37.53502
128	PTE	0.47773	-138.883	71.77869
	IPE	0.29971	-178.411	90.83952
	MRKE	0.58168	-110.138	59.05355
	PTMRKE	0.33789	-176.922	90.24075
256	PTE	0.45859	-217.66	111.6068
	IPE	0.25103	-287.044	145.597
	MRKE	0.55837	-149.181	78.66661
	PTMRKE	0.32598	-263.251	133.7901
512	PTE	0.42731	-399.153	202.6058
	IPE	0.24812	-612.904	308.41
	MRKE	0.54144	-277.416	143.8184
	PTMRKE	0.32587	-480.722	242.8798

Table 15: Root-mean square error, (AIC) criterion and logarithm likelihood criteria for Joe copula when $\tau = 0.3$.

JOE Sample size	Method	RMSE	AIC	LOGL
32	PTE	0.71841	-5.35054	5.12134
	IPE	0.51863	-14.7359	9.38528
	MRKE	1.06945	2.1485	0.04917
	PTMRKE	0.84603	1.55874	2.07229
64	PTE	0.68521	-21.1653	13.1342
	IPE	0.62134	-31.0146	17.99025
	MRKE	1.04992	1.45053	0.54499
	PTMRKE	0.79347	-9.15675	7.51553
128	PTE	0.6379	-42.2662	24.05792
	IPE	0.61594	-92.7233	48.50942
	MRKE	1.02292	-21.328	12.65086
	PTMRKE	0.77322	-40.7035	23.26146
256	PTE	0.72761	-115.09	60.9177
	IPE	0.57132	-249.884	126.8797
	MRKE	0.86669	-59.2224	32.13698
	PTMRKE	0.63472	-118.633	62.54165
512	PTE	0.71878	-232.741	119.9393
	IPE	0.52043	-286.617	146.4281
	MRKE	0.84652	-131.683	69.14675
	PTMRKE	0.63097	-243.648	125.1941

Figures 1,2,3,4 and 5 explain the behavior for all five copula.

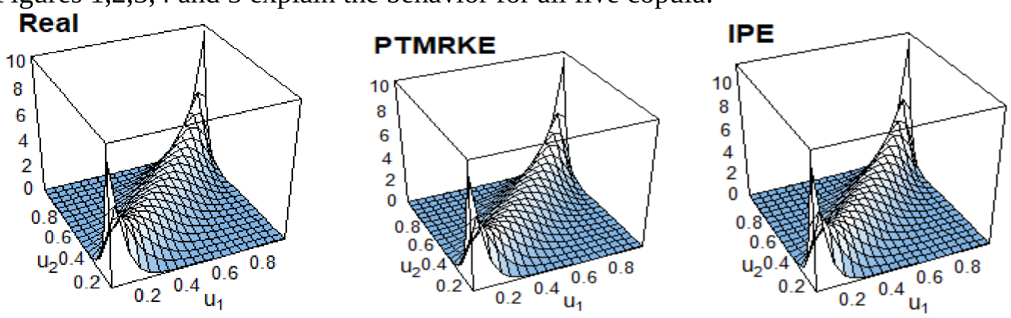

Figure (1) three dimensions Gaussian copula density.

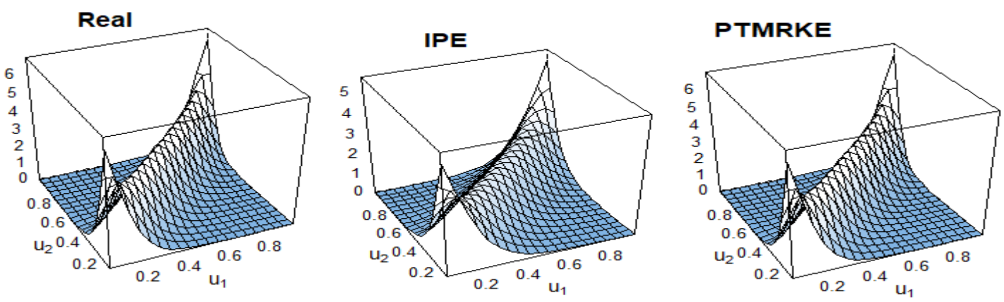

Figure (2) three dimensions for Frank copula density.

P-ISSN 2518-5764
E-ISSN 2227-703X

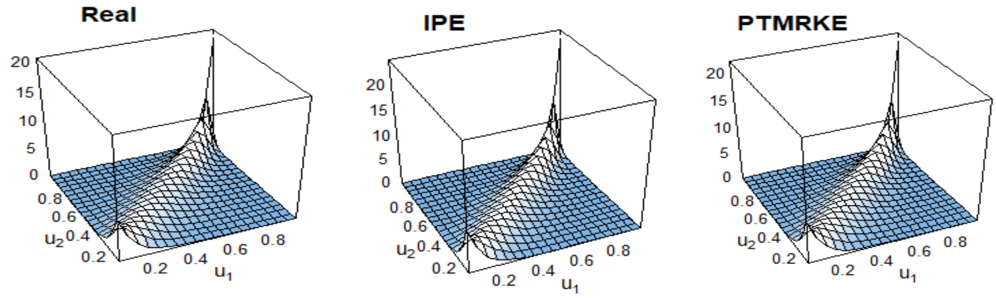

Figure (3) three dimensions for Tawn copula density.

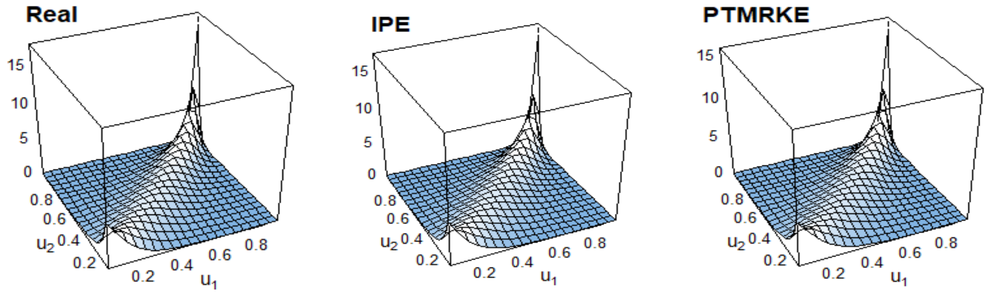

Figure (4) three dimensions for Rotation Tawn copula density.

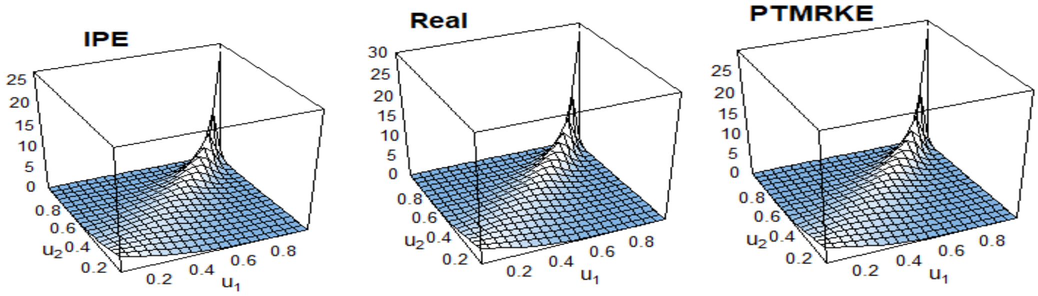

Figure (5) three dimensions for Joe copula density.

4. Conclusions:

Through the results reached in the simulation part of this research, the researcher reached the following conclusions:

1- All the copula functions that have been studied and for all nonparametric estimation methods referred to in the theoretical part and for all sample sizes and at correlation levels, the value of the square root of the mean square error (RMSE) decreases as the sample size increases, while the (LogL) criterion is as maximum as possible, As for the Akaike criteria as minimum as possible.

2- The method of estimating the copula density function using PTMRKE (proposed method) and (IPE) are the two best methods among these methods for the used copulas.

3- The method of nonparametric estimation (IPE) is one of the best methods in estimating the copula density functions due to the fact that the nonparametric function (Gaussian) is more flexible when it is used in choosing the parameter smoothing is fully matrix.

4-The results also indicated that the least-performing estimation method for all values of RMSE and for all sample sizes used is the MRKE method.

5- The proposed method Probit Transform Mirror Reflection Kernel Estimator (PTMRKE) showed handling the boundary bias problem with a probit transform for smoothing observations at boundaries and edges.

6- There is a clear positive effect of the proposed method on the copula functions Tawn, RTawn and Joe; this effect decreases with the large sample size at the copulas (Gaussian, Frank) and in the case of weak and medium dependency.

P-ISSN 2518-5764
E-ISSN 2227-703X

Authors Declaration:
Conflicts of Interest: None
-We Hereby Confirm That All The Figures and Tables In The Manuscript Are Mine and Ours. Besides, The Figures and Images, Which are Not Mine, Have Been Permitted Republication and Attached to The Manuscript.
- Ethical Clearance: The Research Was Approved By The Local Ethical Committee in The University.

References :

1. Alsina, C., Schweizer, B., and Frank, M. J. 2006. Associative functions: triangular norms and copulas. Copyright © by World Scientific Publishing Co.Pte.I.td
2. André, L. M. B. C. M. 2019. Copula models for dependence: comparing classical and Bayesian approaches (Doctoral dissertation). Unversdade Delisboa.
3. Bean, A. T. 2017. Transformations and Bayesian Estimation of Skewed and Heavy-Tailed Densities (Doctoral dissertation, The Ohio State University).
4. Bowman, A. W., and Azzalini, A. 1997. Applied smoothing techniques for data analysis: the kernel approach with S-Plus illustrations. Published in the United States by Oxford University Press INC. New York.
5. Chaloob ,I.H. 2011. Finding Bivariate distribution by using different copulas Function with Application in Biotical Field. PhD thesis, University of Baghdad, Baghdad.
6. Charpentier, A., Fermanian, J. D., and Scaillet, O. 2006. The estimation of copulas: Theory and practice. Copulas: From theory to application in finance, Ensae-Crest and Katholieke Universiteit Leuven; BNP-Paribas and Crest; HEC Genève and Swiss Finance Institute.
7. Cherubini, U., Luciano, E., and Vecchiato, W. 2004. Copula methods in finance. John Wiley & Sons.
8. Chen, L., and Guo, S. 2019. Copulas and its application in hydrology and water resources. Springer Singapore
9. Dawod, E.A.A. 2006. Using the Copula theory for Analyzing the Bivariate survival Function. PhD thesis, University of Baghdad, Baghdad.
10. Dawod, L.A-J. 2022. Structural Reliability Analysis Techniques with Multidimensional Correlation with Application. PhD thesis, University of Baghdad, Baghdad.
11. Deheuvels, P. 1979. La fonction de dépendance empirique etses propriétés. Un test non paramétrique d'indépendance. *Bulletins de l'Académie Royale de Belgique*, 65(1), pp. 274-292.
12. Fabien, F., Perrin, T. T., and Neptune-La Defense, T. 2003. Copula: a new vision for economic capital and application to a four line of Business Company. In Astin Conference.
13. Faugeras, O. P. 2009. A quantile-copula approach to conditional density estimation. *Journal of Multivariate Analysis*, 100(9),pp. 2083-2099.
14. Geenens, G. 2014. Probit transformation for kernel density estimation on the unit interval. *Journal of the American Statistical Association*, 109(505), pp.346-358.
15. Geenens, G., Charpentier, A., and Paindaveine, D. 2017. Probit transformation for nonparametric kernel estimation of the copula density. *Bernoulli*, 23(3), pp.1848-1873.
16. Genest, C., and Favre, A. C. 2007. Everything you always wanted to know about copula modeling but were afraid to ask. *Journal of hydrologic engineering*, 12(4), pp.347-368.
17. Genest, C., Masiello, E., and Tribouley, K. 2009. Estimating copula densities through wavelets. Insurance: *Mathematics and Economics*, 44(20), pp.170-181.
18. Gijbels, I., and Mielniczuk, J. 1990. Estimating the density of a copula function. *Communications in Statistics-Theory and Methods*, 19(2), pp. 445-464.
19. Gramacki, A. 2018. Nonparametric kernel density estimation and its computational aspects (Vol. 37). Cham: Springer International Publishing.
20. Hmood, M.Y. 2005.Comparing Nonparametric Estimators for Probability Density Estimation. PhD thesis, University of Baghdad, Baghdad.

21. Hmood ,M. Y., Abbas, T. M. and Nayef, Q. N. 2008. Nonparametric estimation of A Multivariate Probability Density Function , *Al-Nahrain University Journal* , 11(2), pp.55-63.
22. Hmood, M. Y., and Hamza, Z. F. 2019. On the Estimation of Nonparametric Copula Density Functions. *International Journal of Simulation-Systems, Science & Technology*, 20(2),pp.1-7
23. Loader, C. R. 1996. Local likelihood density estimation. *The Annals of Statistics*, 24(4), pp.1602-1618.
24. Nagler, T. 2014. Kernel methods for vine copula estimation.Master’s Thesis, Technische Universität München Department of Mathematics.
25. Nagler,T. 2021 "Packge Kde vine , Multivariate kernel density Estimation with vine copula " URL https://github.com/tnagler/Kde vine.
26. Nelsen, R. B. 2006. An introduction to copulas. Springer Science & Business Media.
27. Omelka, M., Gijbels, I., and Veraverbeke, N. 2009. Improved kernel estimation of copulas: weak convergence and goodness-of-fit testing. *The Annals of Statistics*, 37(5B), pp. 3023-3058.
28. Patton, A. J. 2002. Applications of copula theory in financial econometrics. Dissertation submitted in partial satisfaction of the requirement for the degree of Doctor of Philosophy in economics. University of California, San Diego.
29. Sklar, M. 1959. Fonctions de répartition à n dimensions et leurs marges. Publ. Inst. Statist. Univ. Paris. 8 ,pp 229–231. MR0125600.
30. Scott D.W. 2009. Multivariate Density Estimation: Theory, Practice, and Visualization, Papers, No.2004,16. New York: John Wiley & Sons.
31. Xue-Kun Song, P. 2000. Multivariate dispersion models generated from Gaussian copula. *Scandinavian Journal of Statistics*, 27(2), pp.305-320.
32. Zeng,X., Ren, J., Sun, M.,Marshall,S., and Durrani,T. 2014. Copulas for statistical signal processing (Part II): Simulation, optimal selection and practical applications. *Signal processing*, 94, pp.681-690.

Journal of Economics and Administrative Sciences
2024; 30 (139), pp. 126-148
P-ISSN 2518-5764
E-ISSN 2227-703X

مقدر كثافة اللب الرابطة بالاعتماد على التحويل الاحتمالي والتحويل الإحتمالي المحسّن

فاطمه هاشم فلحي
جامعة البصرة /كلية الإدارة والاقتصاد/ قسم الإحصاء
البصرة، العراق
fatma.hashem1101a@coadec.uobaghdad.edu.iq
fatima.falhi@uobasrah.edu.iq
http://orcid.org/0009-0008-1170-5635

مناف يوسف حمود
جامعة بغداد / كلية الادارة والاقتصاد/ قسم الاحصاء
بغداد، العراق
munaf.yousif@coadec.uobaghdad.edu.iq
http://orcid.org/0000-0002-1134-9078

Received:1/7/2023 Accepted: 13/8/2023 Published Online First: 29 /2/ 2024

مستخلص البحث:
تستخدم نمذجة الرابطة على نطاق واسع في الإحصائيات الحديثة. اذ تعد مشكلة التحيز الحدودي من المشاكل التي نواجهها عند التقدير بالطرائق اللامعلمية وذلك لان المقدرات اللبية هي الاكثر شيوعا في التقدير اللامعلمي . في هذا البحث تم تقدير دالة كثافة الرابطة باستعمال ثلاثة طرائق لامعلمية من خلال التحويل الاحتمالي للتخلص من مشكلة الانحياز الحدي التي تعاني منها المقدرات اللبية, و باستخدام المحاكاة للطرائق اللامعلمية الثلاثة لتقدير دالة كثافة الرابطة وبالاعتماد على خمسة أنواع من الروابط ذات أحجام عينات مختلفة ومستويات مختلفة من الارتباط بين متغيرات الروابط ولمعلمات مختلفة من تلك الدالة. أظهرت النتائج أن أفضل طريقة هي الجمع بين التحويل الاحتمالي ومقدر انعكاس المرآة اللبي (PTMRKE) عند استخدام جميع دوال الرابطة ولكافة أحجام العينات ، إذا كان الارتباط قويًا تليها طريقة التحويل الاحتمالي المحسنة (IPE). أما في حالة استخدام الارتباطات الضعيفة والمتوسطة ، فقد تبين أن طريقة التحويل الاحتمالي المحسنة (IPE) هي الأفضل ، تليها الطريقة المقترحة (PTMRKE)، اعتمادًا على المعايير(RMSE، LogL، Akaike). وأشارت النتائج أيضًا إلى أن طريقة انعكاس المرآة اللبية تكون ضعيفة عند استخدام الروابط الخمسة.

نوع البحث : ورقة بحثية.
المصطلحات الرئيسة للبحث :دالة الرابطة، التحويل الإحتمالي، دالة الرابطة اللبية، التحويل الإحتمالي المحسن، انعكاس المرآة، التحيز الحدي.

***البحث مستل من اطروحة دكتوراه**