Springer Proceedings in Mathematics & Statistics

R. N. Mohapatra
S. Yugesh
G. Kalpana
C. Kalaivani *Editors*

Mathematical Analysis and Computing

ICMAC 2019, Kalavakkam, India, December 23–24

Springer

Springer Proceedings in Mathematics & Statistics

Volume 344

This book series features volumes composed of selected contributions from workshops and conferences in all areas of current research in mathematics and statistics, including operation research and optimization. In addition to an overall evaluation of the interest, scientific quality, and timeliness of each proposal at the hands of the publisher, individual contributions are all refereed to the high quality standards of leading journals in the field. Thus, this series provides the research community with well-edited, authoritative reports on developments in the most exciting areas of mathematical and statistical research today.

More information about this series at http://www.springer.com/series/10533

R. N. Mohapatra · S. Yugesh · G. Kalpana ·
C. Kalaivani
Editors

Mathematical Analysis and Computing

ICMAC 2019, Kalavakkam, India, December 23–24

Springer

Editors
R. N. Mohapatra
Department of Mathematics
University of Central Florida
Orlando, FL, USA

S. Yugesh
Department of Mathematics
Sri Sivasubramaniya Nadar
College of Engineering
Kalavakkam, Tamil Nadu, India

G. Kalpana
Department of Mathematics
Sri Sivasubramaniya Nadar
College of Engineering
Kalavakkam, Tamil Nadu, India

C. Kalaivani
Department of Mathematics
Sri Sivasubramaniya Nadar
College of Engineering
Kalavakkam, Tamil Nadu, India

ISSN 2194-1009 ISSN 2194-1017 (electronic)
Springer Proceedings in Mathematics & Statistics
ISBN 978-981-33-4645-1 ISBN 978-981-33-4646-8 (eBook)
https://doi.org/10.1007/978-981-33-4646-8

Mathematics Subject Classification: 37C25, 37-XX

This Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd.
The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

Preface: International Conference on Mathematical Analysis and Computing

It is indeed a pleasure and privilege to introduce the proceedings of the selected contributions of the participants of the International Conference on Mathematical Analysis and Computing 2019 (ICMAC 2019) which took place at Sri Sivasubramaniya Nadar College of Engineering, Chennai, India, during December 23–24, 2019.

The main goal of the conference was to provide a platform for experts in mathematical analysis, geometric function theory, and soft computing to meet and interact their results and ideas with each other. Mathematical analysis is one of the most compelling areas of research because of its rich applications. Soft computing is a main pillar of most of the recent research, industrial, and commercial activities. The lectures in the conference were well received by the participants and discussions were fruitful.

The main objective of this "proceedings" is to disseminate recent advances in the studies of diverse areas of mathematical analysis, geometric function theory, and soft computing. These are crystallized in the form of original research articles and expository survey papers.

The papers included in this volume are based on a rigorous peer-review process by the committee of experts in various disciplines. Every submitted paper was first screened by the members of the editorial board, and once it clears the initial screening, it was peer reviewed by at least two potential reviewers in the related area of expertise from the pool of potential reviewers. The paper is accepted if at least two reviewers recommend it for acceptance.

We thank all the invited speakers, reviewers, and the authors who made their valuable contribution toward the success of the conference ICMAC 2019. We would like to thank DST SERB and SSN Trust for the financial support rendered to organize ICMAC 2019. Our sincere gratitude to Mr. Shamim Ahmad, Senior Editor, Springer Nature for his collaboration and timely cooperation.

We express our sincere thanks and gratitude to Dr. Shiv Nadar, Founder, SSN Institutions and Chairman, HCL Technologies; Ms. Roshni Nadar Malhotra, Executive Director and CEO, HCL Corporation, Vice Chairperson, HCL Technologies, Trustee, Shiv Nadar Foundation, Founder and Trustee, The Habitats

Trust; Mr. R. Srinivasan, Director and CEO Redington Ltd. Global Management and Business Leader, Chairman, SSN Institutions; Ms. Kala Vijayakumar, President, SSN Institutions; Members of the Board of Management; Dr. S. Salivahanan, Former Principal, SSN College of Engineering; Dr. V. E. Annamalai, Principal, SSN College of Engineering; Dr. S. Narasimman, COE and Professor, Department of Mathematics, SSN College of Engineering; and Dr. P. Venugopal, Head, Department of Mathematics, SSN College of Engineering for their continuous support and encouragement.

Orlando, USA — Dr. R. N. Mohapatra
Kalavakkam, India — S. Yugesh
Kalavakkam, India — G. Kalpana
Kalavakkam, India — C. Kalaivani

Contents

About the Editors

Prof. R. N. Mohapatra is a Professor of Mathematics at the University of Central Florida, Orlando, USA. He joined the department in August 1984 after spending a year at the York University, Downsview, Ontario, Canada. Prior to that he served as a Professor of Mathematics at the American University of Beirut, Lebanon. He has visited University of Alberta, Edmonton, Canada; Memorial University, Newfound-land, Canada; and the University of Arkansas, Fayetteville as a Visiting Professor to teach teachers during a summer institute supported by NSF. He has published 167 research papers in refereed journals and refereed conference proceedings. He has written two books, one entitled "Biomedical Statistics with Computing" with Dr. Surya Mohapatra and Prof. Mary H. Regier, and the other entitled "Fuzzy Differential Equations" in collaboration with late Prof. V. Lakshmikantham. The two books were published by Research Studies Press and Taylor Francis, respectively. He has also edited eight conference proceedings published by Springer. His areas of current research interest are Approximation Theory, Equilibrium Problems, and Variational Inequalities, Frames in Hilbert C* modules.

Dr. S. Yugesh is Assistant Professor at the Department of Mathematics, Sri Sivasubramaniya Nadar College of Engineering, Tamil Nadu, India. He obtained his Ph.D. degree from Anna University, Chennai, India in 2016. He has written 10 research papers published in peer-reviewed journals, 12 papers in international conference proceedings, and 1 edited volume book. His areas of research are sampling theory, frame theory, splines, and wavelets. He referees articles for professional journals.

Dr. G. Kalpana is Assistant Professor at the Department of Mathematics, Sri Sivasubramaniya Nadar College of Engineering, Tamil Nadu, India. She obtained her Ph.D. degree from Anna University, Chennai, India. She has written some research papers published in peer-reviewed journals and in the national and international conference proceedings. Her current research work includes fuzzy metric spaces and supremum metric on fuzzy sets.

Dr. C. Kalaivani is Assistant Professor at the Department of Mathematics, Sri Sivasubramaniya Nadar College of Engineering, Tamil Nadu, India. She obtained her Ph.D. degree from Alagappa University, Karaikudi, India. She has written some research papers published in peer-reviewed journals and in the national and international conference proceedings. Her current research work includes topology and fuzzy set theory.

A Primality Test and a Theorem on Twin Primes

B. N. Prasad Rao and M. Rangamma

Abstract Any number greater than one can be represented as a product of prime numbers in a unique way, apart from the order in which the prime factors occur. This is known as the Fundamental Theorem of Arithmetic. Thus, prime numbers are solely used to represent prime numbers or composite numbers. There is a difficulty to prove a number as a prime number by using the canonical representation of numbers. We have put a small step in overcoming this difficulty. In this paper, we made a basic and fundamental change in representing positive integers as the sum of a composite number and a prime number. This basic and fundamental change in the representation of numbers helps in testing the primality of numbers through representation of numbers as the sum of a composite number and prime number. By using this representation, we have also proved a theorem on twin primes. Goldbach's conjecture is about the representation of an even number greater than 4 as the sum of two prime numbers. In this paper, we have proved that an odd number under certain conditions get represented as the sum of two prime numbers.

Keywords Primality test · Twin primes · Representation of positive integers · Composite number

Mathematics Subject Classification (2010) 11A41 · 11A51

I thank the Government of India for supporting my research through the University Grants Commission B.S.R.F fellowship.

B. N. Prasad Rao (✉) · M. Rangamma
Department of Mathematics, Osmania University, Hyderabad, Telangana, India
e-mail: prasadrao4@gmail.com

M. Rangamma
e-mail: rangamma1999@gmail.com

R. N. Mohapatra et al. (eds.), *Mathematical Analysis and Computing*,
Springer Proceedings in Mathematics & Statistics 344,
https://doi.org/10.1007/978-981-33-4646-8_1

1 Introduction

A prime number is represented as a product of 1 and itself. A composite number is represented as a product of two or more prime numbers. Thus, prime numbers are used in representing integers. If a number is a prime number, then it will show certain characteristics. For example, we have Wilson's Theorem, which states as "If p is a prime number, then $(p-1)! \equiv -1(modp)$". The converse of this statement is also true. AKS primality test [4] and Miller test [6] uses nature of a prime number in testing the primality of a number. Lucas test [5] and Pepin primality test [7] are based on the form of the given number n. Pocklington primality test [8] requires the number n to satisfy certain conditions before it becomes a prime number. Goldwasser and Kilian primality test [3] is based on Elliptic curves and this idea was modified by Adleman and Huang [1]. Thus, to prove a number as a prime number, we have two choices. The first one is the choice of using the definition of a prime number and the second one is the choice of using the characteristics of a prime number. There are no tests based on the definition of a prime number. Mostly, we prove the number as a prime number by assuming the number as a composite number. Thus, the definition of a prime number as the representation of the product of 1 and itself is not helping in proving the number as a prime number. Thus, our paper aims at proving a number as a prime number by representing the number as the sum of a composite number and prime number.

The unsolved Goldbach's conjecture states that, "every even integer greater than 4 is the sum of two odd prime numbers". Thus, even numbers greater than 4 has a prime number + prime number representation. This unsolved problem suggests a need to study positive integers in terms of sum of two numbers, where either number in the sum is a composite number or prime number or both numbers are prime numbers or both numbers are composite numbers. This paper is the first step in this direction. In Sect. 2 of this paper, we have proved that every integer greater than 5 can be represented as the sum of a composite number and a prime number. Using this representation to natural numbers, we proved in Sect. 4 a criterion for prime number and a criterion for composite number. Clement [2] has given a criterion for twin primes. His criterion is based on Wilson's theorem. In this paper, we have proved a theorem for twin primes, in terms of the representation of a number as the sum of a composite number and a prime number. The theorem on twin prime shows that an odd number under certain conditions get represented as the sum of two prime numbers, which is contrary to Goldbach conjecture. In Sect. 3, we have introduced new concepts in the form of definitions.

2 Representation of Positive Integers

Theorem 1 *Let $n > 5$, $n \in \mathbb{N}$. Then n can be represented as a sum of a composite number and a prime number.*

Proof Let $n > 5, n \in \mathbb{N}$.

Case 1. Let n be any composite number and p be any prime factor of n. Then $n = pb$ where $b \in \mathbb{N}, b > 1$.
When $b = 2$ then $n = 2p$.

$$n = 2(p - 1 + 1)$$

$$n = 2(p - 1) + 2$$

when $b > 2$. Since $n = bp$, and hence

$$n = p(b - 1 + 1)$$

$$= p(b - 1) + p$$

Thus, any composite number n is represented as a sum of a composite number and a prime number.

Case 2. Let n be any prime number and p be any prime number less than n. Then $n - p = 2k$ for some $k \in \mathbb{N}$.
For $k = 1$, we have, $n = p + 2 = (p - 1) + 3 = 2m + 3$ where $m > 1, m \in \mathbb{N}$.
For $k > 1$ we have,

$$n - p = 2k$$

$$n = 2k + p$$

Thus, any prime number n is represented as a sum of a composite number and a prime number.

3 Definitions

Natural numbers are represented as the sum of a composite number and a prime number or sum of two composite numbers or sum of two prime numbers. In view of this, we have the following new concepts in the form of definitions that follow.

Definition 1 Let the representation of a positive integer n be $n = m + p$, where m is a composite number and p a prime number. We call this representation of n as *cp-representation* of n.

Definition 2 Let the representation of a positive integer n be $n = p + q$, where p, q are prime numbers. We call this representation of n as *pp-representation* of n.

Definition 3 Let the representation of a positive integer n be $n = m_1 + m_2$, where m_1, m_2 are composite numbers. We call this representation of n as *cc-representation* of n.

4 Theorem on Primality Test and Twin Primes

Let us consider the following example:

Example 1 Let us consider the natural number $n = 17$. We write various *cp-representations* of the number 17 as follows:

$17 = 15 + 2 = 14 + 3 = 12 + 5 = 10 + 7 = 6 + 11 = 4 + 13$. In each *cp-representation* of 17, we observe that, no prime number divides the corresponding composite number. This motivates us to the following theorem.

Theorem 2 *Let $q > 5$, $q \in \mathbb{N}$. Then q is a prime number if and only if for each cp-representation of q as $q = m + p$, we have the condition $p \nmid m$, where m is a composite number and p a prime number.*

Proof Let $q > 5, q \in \mathbb{N}$ be a prime number.
Let us suppose that, for all *cp-representations* of q as $q = m + p$,where m is a composite number and p a prime number, the condition $p \nmid m$ is false.
Let $q = m + p$ be a *cp-representation* of q such that $p \mid m$.
Then, $q = kp + p$ where $k > 1, k \in \mathbb{N}$.
Thus, $q = (k + 1)p$ and hence q is a composite number.
This contradicts our hypothesis and hence our supposition is false and the theorem must be true.
Conversely. Let $q = m + p$ be any *cp-representation* of q with the condition that $p \nmid m$. We prove that q is a prime number.
Also, let $p_1 < p_2 < \cdots < p_r < q$ be prime numbers less than q.
Suppose, if possible, that q is a composite number, and hence it has some prime factor $p_j \leqslant \sqrt{q}$, where $1 \leqslant j < r$.
$p_j \mid q$, and hence $q = lp_j$, where $l > 1, l \in \mathbb{N}$.
If $l = 2$ then $q = 2(p_j - 1) + 2$. ($p_j \neq 2$, because $q > 5$)
If $l > 2$ then $q = lp_j$.
$q = (l - 1)p_j + p_j$
Thus, q is represented as a sum of a composite and a prime number where the prime number divides the corresponding composite number in the sum. This contradicts our hypothesis.
This completes the proof.

Corollary 1 *Let $q > 8$, $q \in \mathbb{N}$ has a cp-representation $q = m + p$ where m is a composite number, p a prime number and $p \leq \sqrt{q}$. Then q is a prime number if and only if for each cp-representation of q as $q = m + p$, we have the condition $p \nmid m$.*

Corollary 2 *Let $q > 8$, $q \in \mathbb{N}$ has a cp-representation $q = m + p$ where m is a composite number, p a prime number. Then q is a composite number if and only if there exists atleast one cp-representation of q as $q = m + p$ such that $p \mid m$.*

Example 2 Let us consider the odd number 19. The various representations of the number 19 are as follows:

$$19 = 2 + 17 = 3 + 16 = 5 + 14 = 7 + 12 = 9 + 10 = 11 + 8 = 13 + 6 = 15 + 4$$

We note that the number 19 has a *pp-representation*. Also, 19 and 17 are twin prime numbers. When two numbers are not twin primes, then the biggest of the two odd numbers has only *cc-rpresentation* and *cp-representation*. We also note that Goldbach's conjecture states that every even number greater than 4 is the sum of two primes. The same is observed for some odd numbers under certain conditions. We have the following theorem in this regard:

Theorem 3 *Let p, q be prime numbers where $p > q, p > 5$. Then p, q are twin primes if and only if p has a pp-representation.*

Proof Let p, q be twin primes. Then p = 2 + q and hence, the prime number p has a *pp-representation*.
Conversely. Let the representation of a prime number p be *pp-representation*.
Let $p_1 < p_2 < \cdots < p_r < p$ be prime number less than p and $p > q$.
Since given that $q < p$, and hence q must be one of the primes $p_i \in \{p_1, p_2, \ldots, p_r\}$
Since $p > 5$ is an odd prime number, and hence it has a *cp-representation* given by $p = 2^k l + p_i$ for some prime number $p_i < p$, $k \geqslant 1$,$k, l \in \mathbb{N}$.
Since, given that the prime number p has a *pp-representation*, and hence this is possible only when $k = l = 1$ is substituted in above equation
Putting $k = l = 1$ in the equation $p = 2^k l + p_i$, we get $p = 2 + p_r = 2 + q$ where $q = p_r$.
This completes the proof.

Acknowledgements I thank Prof. P. V. Arunachalam, Dr. C. Goverdan, and Prof. V. Darmiah for their encouragement and valuable suggestions. I also thank reviewers for reviewing this article.

References

1. Adleman, L. M., Huang, M.D.: Primality testing and Abelian varieties over finite fields fields. Lecture Notes in Math, Springer, p. 1512 (1992)
2. Clement, P.A.: Congruences for sets of primes. Am. Math. Monthly **56**, 0023–0025 (1949)
3. Goldwasser, S., Kilian, J.: Almost all primes can be quickly certified. Proc. Ann. ACM Symp. Theory Comput. 316–329 (1986)
4. Manindra, A., Neeraj, K., Nitin, S.: Primes is in P. Ann. Math. **160**, 0781–0791 (2004)
5. Michal, K., Florian, L., Lawrence, S.: 17 Lectures on Fermat Numbers: From Number Theory to Geometry, CMS Books in Mathematics, Canadian Mathematical Society, vol. 9, p. 41. Springer (2001)

6. Miller Gary, L.: Riemann's hypothesis and test s for primality. J. Comput. Syst. Sci. **13**, 0300–0317 (1976)
7. Pépin, P.: Sur la formule $2^{2^n} + 1$. Comptes rendus de l'Académie des Sciences de Paris **85**, 0329–0333 (1877)
8. Pocklington H.C.: The determination of the prime or composite nature of large numbers by Fermat's theorem. In: Proceedings of the Cambridge Philosophical Society, vol. 18, pp. 0029–0030

Choosing the Best Copula Function in Mathematical Modeling

Pranesh Kumar and Cigdem Topcu Guloksuz

Abstract Multivariate distributions with any type of marginal distributions can be easily constructed using Sklar's theorem and copula functions. Precisely, a copula function is a parametric representation of a multivariate distribution function which is expressed in terms of marginally uniform random variables on the unit interval. Copula functions possess some appealing mathematical properties, such as they allow scale-free measures of linear/nonlinear stochastic dependence and are useful in simulating families of multivariate distributions. The concept of stochastic tail dependence refers to the clustering of extreme events. Extreme events, for example, in economic systems and in natural hazards contexts generally exhibit tail dependence, and thus it becomes very important to analyze the extreme behavior. Copula functions can measure nonparametric, distribution-free, or scale-invariant nature of dependence and extreme events. Over the past few decades, several examples of copula functions and copula families with one or more real parameters are studied and are applied in various disciplines like statistics, insurance, finance, economics, survival analysis, information theory, image processing, and engineering. In this paper, we aim to discuss copula functions, copula properties, copula families, and simulations using copula functions. Given a large number of copulas available, we will address an important question from the practitioner's points of view as how to choose the most appropriate copula from the family of Archimedean copulas, namely, Clayton, Gumbel, Frank, and Guloksuz–Kumar for fitting the prediction models.

Keywords Copula functions · Multivariate distributions · Extreme events · Dependence · Simulation

P. Kumar (✉)
University of Northern British Columbia, Prince George, BC V2N 4Z9, Canada
e-mail: pranesh.kumar@unbc.ca

C. Topcu Guloksuz
Ankara Science University, 06200 Ankara, Turkey

R. N. Mohapatra et al. (eds.), *Mathematical Analysis and Computing*,
Springer Proceedings in Mathematics & Statistics 344,
https://doi.org/10.1007/978-981-33-4646-8_2

1 Introduction: Copulas and Archimedean Copulas

Sklar [17] introduced copula function and copula in his article in 1959. Copulas describe dependence structure among random variables and have some advantages provided by copula modeling. One of the advantages is constructing the joint probability distribution function independent of its marginal. Further, copula functions can model both symmetric and asymmetric dependence unlike linear correlation measures. For more details about copulas, references are made to Frees and Valdez [4], Matteis [15], Nelsen [16], Genest and Favre [7], Belgorodski [1], Cherubini et al. [2]. In bivariate context, the dependence structure between two random variables is completely described by known bivariate distributions. However, when different types of dependence structures prevail, bivariate copula functions are used to obtain more efficient models.

Precisely, a copula function is a parametric representation of a multivariate probability distribution function which is expressed in terms of marginally uniform random variables on the unit interval. For any pair (X, Y) of continuous random variables, the bivariate probability distribution function (H) can be expressed in the form $H(x, y) = C(u, v), u, v \in (0, 1)$. Equivalently, u and v can be considered as the continuous marginal distributions $F(x)$ and $G(y)$, respectively. Here, C is the copula function with $C : [0, 1]^2 \rightarrow [0, 1]$. Thus, it can be rewritten that $H(x, y) = C(F(x), G(y))$ and that copula function C is a joint probability distribution function with marginal also being probability distribution functions. It should be noted that, if the marginal are continuous, there is a unique copula representation.

The joint probability density function $f(x, y)$, in terms of copula, is expressed by

$$f(x, y) = \Pi_{i=1}^{k} f(x) \times f(y) \times c(F(x), G(y)), \tag{1}$$

where $f(x)$ and $f(y)$ are the marginal density function of X and Y and coupling is provided by the copula density

$$c(u, v) = c(F(x), G(y)) = \partial^2 C(u, v)/\partial u \partial v, \tag{2}$$

if it exists. In case of independent random variables, copula density $c(u, v)$ is identically equal to one. It may be noted that the importance of the above equation $f(x, y)$ is that the independent portion expressed as the product of the marginals can be separated from the function $c(u, v)$ describing the dependence structure or shape. The dependence structure summarized by a copula is invariant under increasing and continuous transformations of the marginals.

A special class of copulas which is being studied extensively is referred to as the Archimedean copulas which can be generated from a function ϕ, called a generator function. For $0 < t < 1$, the generator function ϕ is such that $\phi(1) = 0, \lim_{t \to \infty} (\phi(t)) = \infty, \phi'(t) < 0, \phi''(t) > 0$. Thus, ϕ is a continuous, strictly decreasing, and convex function and always has an inverse, ϕ^{-1}. Additionally, every ϕ that satisfies these conditions can generate a bivariate copula C, known as the Archimedean copula,

$$C(F(x), G(y)) = \phi^{-1}\{\phi(F(x)) + \phi(G(y))\}. \tag{3}$$

It may be noted that the different generator functions result in different Archimedean copulas. The Archimedean copula is indexed by a parameter (θ), called copula parameter. Copula parameter can be real or multidimensional and is embedded in the generator function ϕ. For a given random sample data (X_i, Y_i; $i = 1, 2, \ldots, n$), the copula parameter θ should be estimated to specify the copula function that models the dependence structure. In literature, there exist various techniques to estimate copula parameters [Nelson (16)]. The one-parameter Archimedean copulas have another useful property about the relationship between the copula parameter and nonparametric association measures like the Kendall τ, Spearman ρ. Scale-invariant-dependent measures can be expressed as copula functions of random variables. Two standard nonparametric dependence measures expressed in copula form are as follows:

$$\text{Kendall's tau}: \ \tau = 4 \iint_{I^2} C(u, v)\mathrm{d}C(u, v) - 1 \tag{4}$$

$$\text{Spearman's Rho}: \ \rho = 12 \iint\limits_{I^2} C(u, v)\mathrm{d}u\mathrm{d}v - 3. \tag{5}$$

Kendall's τ can be estimated from the data by

$$\tau = 2 \sum_{i<j} \mathrm{Sign}\big[\big(x_i - x_j\big)\big(y_i - y_j\big)\big]/n(n-1). \tag{6}$$

It may be noted that the linear correlation coefficient is not expressible in a copula form. For Archimedean copulas with generator function ϕ, there is an alternative expression for τ:

$$\tau = 4 \int_0^1 \big((\phi(t))/\big(\phi'(t)\big)\big)dt + 1. \tag{7}$$

The asymptotic properties of this estimator are studied in Kojadinovic and Yan [14]. Apart from rank-based nonparametric strategies, the maximum likelihood methods to estimate the copula parameter are given by Genest and Rivest [6] and Joe [11].

Copula functions, as scale-free measures of linear/nonlinear stochastic dependence, are useful in simulating families of multivariate probability distributions. One of the advantages of working with Archimedean copulas is that an Archimedean copula can be uniquely determined by the function

$$K\phi(t) = t - \big((\phi(t))/\big(\phi'(t)\big)\big), \ \text{for } 0 < t < 1. \tag{8}$$

$K_\phi(t)$ is the distribution function of $C(u, v)$ and a bivariate Archimedean copula is determined by a univariate function $K_\phi(t)$ which is also called Kendall distribution function.

The empirical estimate of copula, $K_n(t)$ from a random sample of size n [16] is given by

$$K_n(t) = \frac{\#(T_i \leq t)}{n+1}, \tag{9}$$

where pseudo-observations $T_i's$ are defined as

$$T_i = H_n(X_i, Y_i) = \frac{\sum_{j=1}^{n} I[\{X_j \leq X_i\, Y_j \leq Y_i\}]}{\text{n + 1}}, i = 1, 2, \ldots n \tag{10}$$

To select the suitable Archimedean copula function $K_\phi(t)$ from the family of Archimedean copulas, Frees and Valdez [4] considered the degree of closeness between $K_n(t)$ and $K_\phi(t)$ and suggested minimum distance measure

$$MD = \int \left[K_n(t) - K_\phi(t)\right]^2 dK_n(t). \tag{11}$$

Other copula selection criteria are based on information measures such as Akaike information criteria (AIC) and Bayesian information criteria (BIC), both being dependent on likelihood functions and are

$$AIC = -2\sum_{i=1}^{n} \ln\left[c\left(u_{i,1}, u_{i,2}\right); \theta\right] + 2k, \tag{12}$$

$$BIC = -2\sum_{i=1}^{n} \ln\left[c\left(u_{i,1}, u_{i,2}\right); \theta\right] + \ln(nk). \tag{13}$$

The preferred copula is with the lowest AIC/BIC values. Also, some derivations of AIC and other information criteria for model selection are given in Grønneberg and Hjort [8].

What follows now, the bivariate Archimedean copulas studied in this paper are given in Table 1. The expressions of Kendall's τ in terms of copula parameters are provided in the last column. The copula parameter in each case measures the degree of dependence and controls the association between two variables. In Table 2, the expressions of Archimedean copulas, their generator functions, and distribution functions are given. As an example of the graphic representation of copula, the scatterplot of the Guloksuz–Kumar (NewCop) copula [9] and its distribution function are shown in Fig. 1.

Table 1 Archimedean copulas and Kendall's τ

Family	Bivariate copula $C(u,v), 0<u,v<1$	Copula parameter space	τ
Clayton	$\left(u^{-\theta}+v^{-\theta}-1\right)^{-1/\theta}$	$\theta>1$	$\frac{\theta}{\theta+2}$
Gumbel	$\exp\left[-\left[(-\ln u)^{\theta}+(-\ln v)^{\theta}\right]^{1/\theta}\right]$	$\theta\geq 1$	$\frac{\theta-1}{\theta}$
Frank	$\left(-\frac{1}{\theta}\right)\ln\left\{\frac{(1-e^{-\theta})-(1-e^{-\theta u})(1-e^{-\theta v})}{(1-e^{-\theta})}\right.$	$-\infty<\theta<\infty$	$1-\frac{4}{\theta}\left[1-D_1^*(\theta)\right]$
Guloksuz–Kumar	$1-\frac{1}{\theta}\ln\left(e^{\theta(1-u)}+e^{\theta(1-v)}-1\right)$	$\theta>0$	$\frac{4\left(1-e^{-\theta}-\theta\right)}{\theta^2}+1+$ 1

$D_n^*(\theta)=n/\theta^n\int_0^{\theta}\frac{t^n}{(e^t-1)dt)}$, $n>0$ is a Debye function

Table 2 Archimedean copulas, generator functions, and Copula distribution function

Family	Generator $\phi(t), 0<t<1$	Distribution function $K_{\phi}(t)=t-\frac{\phi(t)}{\phi'(t)}, 0<t<$ 1
Clayton	$t^{-\theta}-1$	$t-\frac{t^{\theta+1}-t}{\theta}$
Gumbel	$(-\ln t)^{\theta}$	$t-\frac{t\ln t}{\theta}$
Frank	$-\ln\frac{e^{-\theta t}-1}{e^{-\theta}-1}$	$t-\frac{\ln\frac{e^{-\theta t}-1}{e^{-\theta}-1}}{\theta}\left(e^{\theta t}-1\right)$
Guloksuz–Kumar	$e^{\theta(1-t)}-1$	$t-\frac{e^{\theta(1-t)}-1}{-\theta e^{-\theta(t-1)}}$

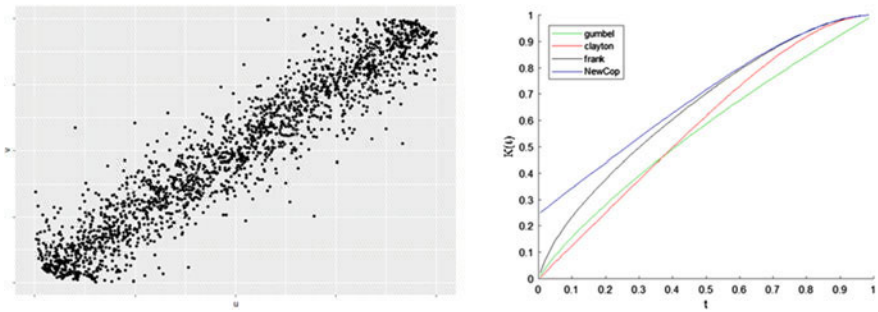

Fig. 1 Scatterplot and distribution function of the Guloksuz–Kumar (NewCop) copula

2 Copula Tail Dependence Measures

The concept of stochastic tail dependence refers to the clustering of extreme events. Extreme events, for example, in economic systems and in natural hazards contexts generally exhibit tail dependence, and thus it becomes very important to analyze the extreme behavior. The definition of tail dependence is the limiting probability that a random variable exceeds a certain threshold, given that another random variable already exceeds that threshold. Copula functions can measure nonparametric, distribution-free, or scale-invariant nature of dependence and extreme events. Joe [11] defines the upper tail dependence:

$$\lambda_U := \lim_{u \to 1}[\{1 - 2u + C(u, u)\}/(1 - u)] \tag{14}$$

$C(u, v)$ has upper tail dependence for $\lambda_U \in (0, 1]$ and no upper tail dependence for $\lambda_U = 0$.

Similarly, lower tail dependence in terms of copula is defined as follows:

$$\lambda_L := \lim_{u \to 0}[C(u, u)/u], \tag{15}$$

and copula $C(u, v)$ has lower tail dependence for $\lambda_L \in (0, 1]$ and no lower tail dependence for $\lambda_L = 0$. This measure is extensively used in extreme value theory. It is the probability that one variable is extreme given that other is extreme. Tail measures are copula based and copula is related to the full distribution via quantile transformations, i.e.,

$$C(u, v) = F\left(F_1^{-1}(u), F_2^{-1}(v)\right), \tag{16}$$

for all $u, v \in (0, 1]$ in the bivariate case.

3 Copula Simulation

The following steps are used to simulate data from a bivariate copula. The procedure mainly consists of first to estimate the marginal distributions and then to specify the copula function. Marginal distributions can be estimated by empirical or parametric ways. For simulation, let (U, V) be a random pair from a bivariate Archimedean copula, $\phi(t)$ the generator function, and $K(t)$ the distribution function of copula. Then, n-pairs of data $(X_i, Y_i; i = 1, 2, \ldots, n)$ from a bivariate Archimedean copula can be simulated from the following steps:

(1) Draw two independent random numbers, p and q, from Uniform (0, 1).
(2) $t = K^{-1}(q)$.
(3) $u = \phi^{-1}[p\phi(t)]$ and $v = \phi^{-1}[(1 - p)\phi(t)]$.
(4) $X = F^{-1}(u)$ and $Y = F^{-1}(v)$.
(5) Repeat the above steps n times to get n pairs of data $(X_i, Y_i; i = 1, 2, \ldots, n)$.

It should be also noted that for the step 2, when the inverse of $K(t)$ does not have the closed form, the following equation can be solved by numerical methods, like the Newton–Raphson method:

$$\left|t - \left((\phi(t))/\left(\phi'(t)\right)\right)\right| - q = 0. \tag{17}$$

4 Empirical Study

Several applications of copula functions and copula families with one or more real parameters are studied in various disciplines like statistics, insurance, finance, economics, survival analysis, information theory, image processing, and engineering. We now address an important concern from the practitioner's point of view as how to choose the most appropriate copula as a model from the large family of Archimedean copulas for a given sample dataset.

4.1 Modeling and Simulation

For this objective, we begin with fitting the linear model for making predictions and analyzing the prediction errors for choosing the best copula among the Archimedean copulas, namely, Clayton [3], Gumbel [10], Frank [5], and Guloksuz–Kumar copulas [9]. For more details on copula-based prediction, refer to Kumar and Shoukri [12, 13]. The calculation steps are as follows:

i. Randomly generate 1000 pairs of (X, Y) data from the bivariate standard normal distribution and for the Pearson correlation coefficient $=$ 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.75, 0.8, 0.85, 0.9, 0.95, 0.99.
ii. Calculate Kendall's τ and the copula parameter θ using the expressions in the last column of Table 1.
iii. For each copula, simulate 1000 pairs of (X, Y) and repeat simulations 100 times.
iv. For the simulated data in the previous step (iv), obtain least squares estimates a and b to fit the linear regression model to predict Y from X, i.e., $E(Y) = \alpha + \beta X + \varepsilon$.
v. The estimated intercept and slope coefficient values (a and b), of the parameters (α and β), for the copula-based models are, respectively, the averages of the 100 (a and b) values from the fitted models in step (iv).
vi. Calculate the 95% confidence intervals (CIs) of the slope parameter and the mean absolute prediction error ($MAPE$)

$$MAPE = \sum_{i=1}^{n} 100\left|\frac{y_i - \hat{y}_i}{ny_i}\right| \tag{18}$$

In Table 3, values of the Pearson correlation coefficient r, Kendall τ, and copula parameters θ of Clayton, Gumbel, Frank, and Guloksuz–Kumar copulas are provided.

It is noted that for the range of values of the correlation coefficient r between 0.1 and 0.99, the range of values of Kendall τ is between 0.0852 and 0.9132 and the range of copula parameters are Clayton (0.1864, 21.0397), Gumbel (1.0932, 11.5198), Frank (0.717, 44.3712), and Guloksuz–Kumar (2.9798, 45.0566). Since copula parameter is a measure of the dependence or association between two variables like Pearson correlation coefficient r and Kendall τ, the range of Frank copula parameter (43.5995) is the maximum, followed by the range of Guloksuz–Kumar copula parameter (42.0768) and is the minimum (10.4266) for the Gumbel copula parameter.

4.2 *Choosing Copula for the Prediction Modeling Using Mean Absolute Prediction Error (MAPE)*

We summarize fitted prediction models, 95% confidence intervals (CI) of regression coefficients, and mean absolute prediction error ($MAPE$) of the models in Table 4. It is clearly evident from the minimum distance measure ($MAPE$) values that the Guloksuz–Kumar-copula-based prediction model has the minimum prediction error for all the Pearson correlation coefficient (r) considered except $r = 0.1$, where the Clayton-copula-based prediction model has the minimum prediction error (Table 4).

Table 3 Kendall τ and copula parameters θ

Correlation r	Kendall τ	Clayton θ	Gumbel θ	Frank θ	Guloksuz–Kumar θ
0.1	0.0852	0.1864	1.0932	0.7717	2.9798
0.2	0.1199	0.2724	1.1362	1.0916	3.1721
0.3	0.1726	0.4172	1.2086	1.5920	3.4921
0.4	0.2319	0.6038	1.3019	2.1834	3.8991
0.5	0.3415	1.0373	1.5186	3.4046	4.8259
0.6	0.4190	1.4426	1.7213	4.4295	5.6764
0.7	0.4940	1.9524	1.9762	5.6278	6.7319
0.75	0.5252	2.2126	2.1063	6.2148	7.2666
0.8	0.5810	2.7732	2.3866	7.4430	8.4118
0.85	0.6587	3.8592	2.9296	9.7401	10.6144
0.9	0.7037	4.7493	3.3746	11.5815	12.4109
0.95	0.8060	8.3070	5.1535	18.8115	19.5602
0.99	0.9132	21.0397	11.5198	44.3712	45.0566

Table 4 Models for prediction and mean absolute prediction error (MAPE)

Correlation	Data	Intercept	Slope (std. error)	95% CI (slope)	MAPE
$r = 0.1$	BiNormal (0, 1)	−0.0444	0.1270 (0.0318)	0.0646–0.1895	1.3121
	Clayton	−0.0033	0.1380 (0.0312)	0.0768–0.1992	1.2090
	Gumbel	−0.0012	0.1399 (0.0311)	0.0789–0.2008	1.2109
	Frank	−0.0015	0.1218 (0.0314)	0.0602–0.1833	1.3096
	Guloksuz–Kumar	0.1246	−0.0952 (0.0315)	−0.1569–0.0336	1.5400
$r = 0.2$	BiNormal (0, 1)	0.0300	0.1824 (0.0316)	0.1203–0.2445	1.7407
	Clayton	0.001	0.1925 (0.0311)	0.1315–0.2534	1.7874
	Gumbel	0.0020	0.1961 (0.0312)	0.1349–0.2573	1.8037
	Frank	0.0035	0.1702 (0.0310)	0.1095–0.2310	1.6862
	Guloksuz–Kumar	0.1029	−0.0383 (0.0315)	−0.1000–0.0234	1.3838
$r = 0.3$	BiNormal (0, 1)	−0.0194	0.2677 (0.0305)	0.2079–0.3275	2.0681
	Clayton	0.0035	0.2778 (0.0305)	0.2180–0.3374	2.1367
	Gumbel	0.0035	0.2801 (0.0306)	0.2201–0.3399	2.1475
	Frank	0.0026	0.2443 (0.0307)	0.1840–0.3045	1.9797
	Guloksuz–Kumar	0.0813	0.0471 (0.0315)	−0.0146–0.1089	1.3641
$r = 0.4$	BiNormal (0, 1)	0.0286	0.3636 (0.0284)	0.3078–0.4193	1.8016
	Clayton	0.0010	0.3574 (0.0295)	0.2996–0.4153	1.7912
	Gumbel	0.0047	0.3655 (0.0295)	0.3077–0.4233	1.8125
	Frank	0.0073	0.3209 (0.0299)	0.2622–0.3797	1.6782
	Guloksuz–Kumar	0.0578	0.1405 (0.0313)	0.0791–0.2019	1.2514
$r = 0.5$	BiNormal (0, 1)	0.0324	0.4706 (0.0261)	0.4195–0.5217	2.8881
	Clayton	0.0074	0.5017 (0.0275)	0.4479–0.5556	3.0339
	Gumbel	0.0013	0.5037 (0.0272)	0.4504–0.5569	3.0441
	Frank	0.0037	0.4708 (0.0280)	0.4159–0.5257	2.8890
	Guloksuz–Kumar	0.0290	0.3395 (0.0298)	0.2811–0.3979	2.2896
$r = 0.6$	BiNormal (0, 1)	0.0548	0.5927 (0.0242)	0.5453–0.6401	2.1512
	Clayton	0.0023	0.6031 (0.0254)	0.5533–0.6529	2.1413
	Gumbel	0.0004	0.6064 (0.0251)	0.5572–0.6556	2.1488
	Frank	0.0016	0.5702 (0.0263)	0.5187–0.6216	2.0580
	Guloksuz–Kumar	0.0144	0.4665 (0.0280)	0.4117–0.5213	1.8102
$r = 0.7$	BiNormal (0, 1)	−0.0081	0.7132 (0.0227)	0.6687–0.7577	2.9027
	Clayton	−0.0015	0.6773 (0.0233)	0.6316–0.7229	2.7072
	Gumbel	−0.0006	0.6936 (0.0227)	0.6491–0.7380	2.7499
	Frank	−0.0002	0.6522 (0.0241)	0.6050–0.6994	2.6106
	Guloksuz–Kumar	0.0072	0.5814 (0.0256)	0.5312–0.6316	2.2993
$r = 0.75$	BiNormal (0, 1)	−0.0272	0.7293 (0.0210)	0.6880–0.7706	3.9980

(continued)

Table 4 (continued)

Correlation	Data	Intercept	Slope (std. error)	95% CI (slope)	MAPE
	Clayton	−0.0003	0.7065 (0.0223)	0.6629–0.7502	3.9776
	Gumbel	−0.0015	0.7192 (0.0218)	0.6764–0.7620	4.0356
	Frank	−0.0021	0.6870 (0.0231)	0.6417–0.7322	3.8765
$r = 0.8$	BiNormal (0, 1)	0.0140	0.7801 (0.0189)	0.7430–0.8173	2.6190
	Clayton	0.0012	0.7634 (0.0205)	0.7231–0.8036	2.5708
	Gumbel	0.0009	0.7830 (0.0196)	0.7445–0.8214	2.6253
	Frank	0.0034	0.7479 (0.0212)	0.7064–0.7895	2.5276
	Guloksuz–Kumar	0.0044	0.7033 (0.0224)	0.6593–0.7472	2.4066
$r = 0.85$	BiNormal (0, 1)	0.00001	0.8514 (0.0162)	0.8197–0.8831	2.6908
	Clayton	0.0009	0.8217 (0.0180)	0.7864–0.8570	2.6088
	Gumbel	0.0003	0.8498 (0.0166)	0.8173–0.8824	2.6861
	Frank	0.0003	0.8138 (0.0185)	0.7776–0.8499	2.5881
	Guloksuz–Kumar	0.0029	0.7942 (0.0195)	0.7559–0.8324	2.5334
$r = 0.9$	BiNormal (0, 1)	0.0146	0.9014 (0.0142)	0.8736–0.9292	1.4276
	Clayton	0.0016	0.8576 (0.0164)	0.8255–0.8896	1.3725
	Gumbel	0.0011	0.8833 (0.0146)	0.8546–0.9120	1.4035
	Frank	0.0010	0.8469 (0.0167)	0.8142–0.8797	1.3613
	Guloksuz–Kumar	0.0007	0.8339 (0.0175)	0.7997–0.8681	1.3468
$r = 0.95$	BiNormal (0, 1)	0.0056	0.9654 (0.0096)	0.9466–0.9841	2.0714
	Clayton	0.0016	0.9199 (0.0126)	0.8953–0.9445	1.9683
	Gumbel	0.0001	0.9491 (0.0100)	0.9294–0.9687	2.0163
	Frank	0.0001	0.9191 (0.0125)	0.8946–0.9436	1.9618
	Guloksuz–Kumar	0.0015	0.9172 (0.0128)	0.8922–0.9422	1.9603
$r = 0.99$	BiNormal (0, 1)	−0.0027	0.9899 (0.0043)	0.9815–0.9983	0.9086
	Clayton	−0.0008	0.9717 (0.0074)	0.9573–0.9861	0.8891
	Gumbel	−0.0008	0.9898 (0.0046)	0.9807–0.9989	0.9024
	Frank	0.8405	2.1743 (0.0460)	2.0842–2.2644	5.7300
	Guloksuz–Kumar	−0.0010	0.9726 (0.0072)	0.9585–0.9867	0.8929

5 Concluding Remarks

Copula functions and copula families with one or more real parameters have become very popular recently for the modeling and simulation in various disciplines like statistics, insurance, finance, economics, survival analysis, information theory, image processing, and engineering. Since there are a large number of copulas available, we investigated an important question as how to choose the most appropriate copula from the family of Archimedean copulas, namely, Clayton, Gumbel, Frank, and Guloksuz–Kumar for making the predictions from the linear models.

We have focused at the bivariate normal population, $N(0, 1)$, since most of the application data do follow the exact or approximate multivariate normal distributions and data can easily be standardized. Further, we have considered the bivariate normal populations with the Pearson correlation coefficient $r = 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.75, 0.8, 0.85, 0.9, 0.95, 0.99$. The Guloksuz–Kumar-copula-based prediction models have attained the minimum prediction error for all the Pearson correlation coefficient (r) values considered, except $r = 0.1$, where the Clayton-copula-based prediction model resulted in the minimum prediction error. However, it may further be noted that when response and explanatory variable have a correlation of very small magnitude like $r = 0.1$, it implies that there is almost no association and the explanatory variable is of no use in making predictions about the response variable, and hence no need for fitting the model.

References

1. Belgorodski, N.: Selecting pair copula families for regular vines with application to the multivariate analysis of European stock market indices. Master Thesis, Technische Universitt Munchen (2010)
2. Cherubini, U., Luciano, E.E., Vecchiato, W.: Copula Methods in Finance. John Wiley, NY (2004)
3. Clayton, D.G.: A Model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence. Biometrika **65** (141–151 (1978)
4. Frees, E.W., Valdez, E.A.: Understanding relationships using copulas. North Am. Actuar. J. **2**, 1–25 (1998)
5. Frank, M.J.: On the simultaneous associativity of F(x, y) and x + y-F(x, y). Aequ. Math. **19**, 194–226 (1979)
6. Genest, C., Rivest, L.P.: Statistical inference procedures for bivariate Archimedean copulas. Am. Stat. **88**(423), 1034–1043 (1993)
7. Genest, C., Favre, A.: Everything you always wanted to know about copula modeling but were afraid to ask. J. Hydrol. Eng. **12**, 347–368 (2007)
8. Grønneberg, S., Hjort, N.L.: The copula information criteria. Scand. J. Stat. **41**, 436–459 (2014)
9. Guloksuz, C.T., Kumar, P.: On a new bivariate one parameter Archimedean copula function and its application to modeling dependence in climate and life sciences. Stat. Appl. **17**(2), 23–36 (2019)
10. Gumbel, E.J.: Bivariate exponential distributions. J. Am. Stat. Assoc. **55**, 698–707 (1960)
11. Joe, H.: Multivariate Models and Dependence Concepts, 1st edn. Chapman Hall Ltd, U.K (1997)
12. Kumar, P., Shoukri, M.M.: Aortic stenosis using the Archimedean copula methodology. J. Data Sci. **6**, 173–187 (2008)
13. Kumar, P., Shoukri, M.M.: Copula based prediction models: an application to aortic regurgitation study. BMC Med. Res. Methodol. **7**(21), 1–9 (2007)
14. Kojadinovic, I., Yan, J.: Comparison of three semiparametric methods for estimating dependence parameters in copula models. Insur. Math. Econ. **47**, 52–63 (2010)
15. Matteis De, R.: Fitting copulas to data. Diploma Thesis, Institute of Mathematics of University Zurich (2001)
16. Nelsen, R.: An Introduction to Copulas, 2nd edn. Springer, New York (2006)
17. Sklar, A.: Fonctions de repartition a *n* dimensions et leurs marges. Publications del Institut de Statistique del Universite de Paris **8**, 229–231 (1959)