SELÇUK UNIVERSITY

IWCMA

CMA
CONSTRUCTIVE MATHEMATICAL ANALYSIS

2023
Proceedings of International Workshop: Constructive Mathematical Analysis

06-08 July 2023 Konya-TÜRKİYE

2023 Proceedings of International Workshop: Constructive Mathematical Analysis

ISBN: 978-625-00-1216-1

July 6-8, 2023
Konya, Türkiye

CMA
CONSTRUCTIVE MATHEMATICAL ANALYSIS

Message from Editor

Dear Participants,

It's my pleasure to edit the Proceedings Book of "2nd International Workshop: Constructive Mathematical Analysis". Our workshop is an activity of the journal Constructive Mathematical Analysis and supported by Scientific Research Projects Coordinatorship of Selçuk University. By organizing this workshop, our main aim was to promote, encourage, and provide a forum for the academic exchange of ideas and recent research works. The workshop presents new results and future challenges, in a series of virtual keynote lectures and virtual contributed short talks. In our workshop, we provide a forum for mathematicians to communicate recent research results in the areas of Real analysis, Complex analysis, Potential theory, Special functions and their applications, Matrix analysis, Approximations and expansions, Harmonic analysis, Fourier analysis, Integral transforms-operational calculus, Functional analysis, Fixed Point Theory, Operator theory, Miscellaneous applications of functional analysis, Convex and geometric analysis, Stochastic analysis, Numerical analysis, and the Applications of these fields in other areas. The conference was face to face, and only one presentation was allowed for each participant. The language of all presentations was English, and submissions were peer-reviewed by at least two referees. This proceedings book also includes the papers which are peer-reviewed by at least two referees.
Thanks.

Assoc. Prof. Dr. Tuncer Acar
Selçuk University
Editor of Proceedings Book of ICOMSS'22

Chairmen of Workshop

Tuncer Acar
Selçuk University, Türkiye
tunceracar@ymail.com

Osman Alagöz
Bilecik Şeyh Edebali University, Bilecik, Türkiye
osman.alagoz@bilecik.edu.tr

Organizing Committee

Özlem Acar
Selçuk University, Türkiye
acarozlem@ymail.com

Borislav R. Draganov
Sofia University "St. Kliment Ohridski", Bulgaria
bdraganov@fmi.uni-sofia.bg

Peter Massopust
Technical University of Munich, Germany
massopust@ma.tum.de

Fırat Özsaraç
Kırıkkale University, Türkiye
firatozsarac@kku.edu.tr

Metin Turgay
Selçuk University, Türkiye
metinturgay@yahoo.com

Ioan Raşa
Technical University of Cluj-Napoca, Romania
Ioan.Rasa@math.utcluj.ro

Yoshihiro Sawano
Chuo University, Japan
yoshihiro-sawano@celery.ocn.ne.jp

Scientific Committee

Francesco Altomare
University of Bari Aldo Moro, Italy

Ali Aral
Kırıkkale University, Türkiye

Raul Curto
University of Iowa, USA

Harun Karslı
Abant Izzet Baysal University, Türkiye

Kehe Zhu
State University of New York, USA

Ferenc Weisz
Eotvos University, Hungary

Gianluca Vinti
University of Perugia, Italy

Secretary

Abdulkadir Eke
Selçuk University, Türkiye
abdulkadir-eke42@hotmail.com

Eşref Erdoğan
Selçuk University, Türkiye
esreferdogan@yahoo.com

Sadettin Kurşun
Selçuk University, Türkiye
sadettinkursun@yahoo.com

Dilek Özer
Selçuk University, Türkiye
dilekozer@yahoo.com

Aybala Sevde Özkapu
Selçuk University, Türkiye
aybalasevdeozkapu@yahoo.com

Contents

Some properties of squared Chlodovsky operators

Hüseyin Erhan Altın*[1] and Harun Karslı[2]

ABSTRACT. The main purpose of this paper is to define a sequence of positive linear operators by means of the squared Chlodovsky basis functions. As a consequence of some certain inequalities we state that the second central moments of the constructed operators are smaller than the corresponding ones of the classical Bernstein-Chlodovsky operators. Furthermore, we estimate the rate of convergence in terms of the modulus of continuity and the class of Lipschitz functions.

2020 MATHEMATICS SUBJECT CLASSIFICATIONS: 41A25, 41A35, 41A36

KEYWORDS: Chlodovsky operators, Linear positive operators, Rate of convergence

1. INTRODUCTION

The basis of the theory of approximation is the theorem discovered by Weierstrass in 1985 and the first constructive proof of this theorem was given by Bernstein [3] in 1912. He introduced a sequence of polynomials $B_m : C[0,1] \to C[0,1]$ defined by

$$(B_m f)(x) = \sum_{k=0}^{m} f\left(\frac{k}{m}\right) p_{m,k}(x) \;,\quad x \in [0,1],$$

where Bernstein basis function is given by

$$p_{m,k}(x) = \binom{m}{k} x^k (1-x)^{m-k} \;,\; m \in \mathbb{N}. \tag{1}$$

Later it was discovered that Bernstein polynomials have numerous significants properties, so new applications and generalizations are being found of it. The aim of these generalizations is to provide appropriate and powerful tools to application areas. One of these generalizations is the classical Bernstein-Chlodovsky operators are defined for a function f defined on $[0,\infty)$ and bounded on every finite interval $[0,b] \subset [0,\infty)$ by

$$(C_m f)(x) = \sum_{k=0}^{m} f\left(\frac{b_m}{m}k\right) p_{m,k}\left(\frac{x}{b_m}\right) \tag{2}$$

where $p_{m,k}$ denotes as usual as (1) and $(b_m)_{m=1}^{\infty}$ is a positive increasing sequence of reals with the properties

$$\lim_{m\to\infty} b_m = \infty \;,\; \lim_{m\to\infty} \frac{b_m}{m} = 0. \tag{3}$$

These operators were introduced by Chlodovsky [4] in 1937 as a generalization of the Bernstein polynomials $(B_m f)(x)$ on an infinete interval. Although there are many works on these operators, readers can be find some results, collectively, on Chlodovsky operators in [9].

We can say that the last century has been very productive by means of working with linear positive operators. However working on squared type operators has become popular at the last decade. Now we give a look some remarkable studies on this topic. In [5], the authors obtained a new representation of the sum of the squared Bernstein polynomials and used it to validate a conjecture asserting that this sum is a convex function. They extended the result to some other classical approximation operators; Mirakjan-Favard-Szăsz operators, Meyer-König and Zeller operators and King-type operators. In [6], the authors provided some estimates of the second central moment of the squared Bernstein polynomials and also estimated the rate of approximation in terms of the modulus of continuity. In [7] and [8], the author obtained the Voronovskaya formula for sequence of positive linear operators constructed using the squared Bernstein polynomials and studied some aproximation properties of positive linear operators defined by means of the powered Baskakov basis, respectively. In

[1] and [2], the author(s) derived a complete asymptotic expansion for a sequence of positive linear approximation operators defined by means of the squared Bernstein basis polynomials, Favard-Szăsz-Mirakjan fundamental functions and Baskakov fundamental functions. Also they studied the asymptotic properties of operators defined by means of squared Meyer-König and Zeller fundamental functions. In [10], the authors introduced a family of neural networks of multivatiate square rational Bernstein operators defined by extending the artificial neural networks multivariate Bernstein by using square Bernstein polynomials and they studied the behavior of this neural network. And in [11], the authors defined a sequence of positive linear operators by means of the squared Szăsz-Mirakjan basis functions and estimated the rate of convergence in terms of the modulus of continuity and the class of Lipschitz functions. Furthermore, they showed the comparison and convergence of these operators with the help of some illustrative graphics.

In the light of these works, we consider the rational functions,

$$c_{m,k} = \frac{p_{m,k}^2}{\sum\limits_{i=0}^{m} p_{m,i}^2}\ ,\ k = 0, 1, ..., m, \tag{4}$$

and define the positive linear Bernstein-Chlodovsky type rational operators $(SC_m f)$ by

$$(SC_m f)(x) = \sum_{k=0}^{m} f\left(\frac{b_m}{m}k\right) c_{m,k}\left(\frac{x}{b_m}\right) \tag{5}$$

where $c_{m,k}$ defined by (4) and $(b_m)_{m=1}^{\infty}$ is a positive increasing sequence of reals with the properties (3). This work is organized as follows. In the next section, we give some auxiliary results. And in the final section, we prove the main results which are related with the second central moments and rate of convergence of these operators.

2. Auxiliary results

In order to prove the main results, the following function will be the essential tool. For $m \in \mathbb{N}$ and $u \in [0, 1]$ define

$$g_m(u) = \frac{\int\limits_0^1 \frac{ut(1-ut)^{m-1}}{\sqrt{t(1-t)}}dt}{\int\limits_0^1 \frac{(1-ut)^m}{\sqrt{t(1-t)}}dt}. \tag{6}$$

Integration by parts shows that the integrals in (6) are differentiable with respect to the parameter u. Also define

$$M_m = \sup_{u\in[0,1]} g_m(u)\ ,\ m = 1, 2, ... \tag{7}$$

Lemma 2.1. *The sum of squared Bernstein-Chlodovsky basis function satisfies the following equality*

$$\sum_{k=0}^{m} p_{m,k}^2\left(\frac{x}{b_m}\right) = \frac{1}{\pi}\int\limits_0^1 \frac{\left(1 - 4\frac{x}{b_m}\left(1 - \frac{x}{b_m}\right)t\right)^m}{\sqrt{t(1-t)}}dt$$

for $x \in [0, b_m]$ and $m \in \mathbb{N}$.

Proof. Since

$$\sum_{k=0}^{m} p_{m,k}\left(\frac{x}{b_m}\right)e^{ik\theta} = \left(\frac{x}{b_m}e^{i\theta} + 1 - \frac{x}{b_m}\right)^m,$$

by Parseval's identity, we have

$$\begin{aligned}\sum_{k=0}^{m} p_{m,k}^2\left(\frac{x}{b_m}\right) &= \frac{1}{2\pi}\int\limits_{-\pi}^{\pi}\left|\frac{x}{b_m}e^{i\theta} + 1 - \frac{x}{b_m}\right|^{2m} d\theta \\ &= \frac{2}{\pi}\int\limits_0^{\pi/2}\left(1 - 4\frac{x}{b_m}\left(1 - \frac{x}{b_m}\right)\sin^2\theta\right)^m d\theta,\end{aligned}$$

and with $\sin^2\theta = t$, we have

$$\sum_{k=0}^{m} p_{m,k}^2\left(\frac{x}{b_m}\right) = \frac{1}{\pi}\int_0^1 \frac{\left(1-4\frac{x}{b_m}\left(1-\frac{x}{b_m}\right)t\right)^m}{\sqrt{t\left(1-t\right)}}dt. \tag{8}$$

□

Lemma 2.2. *The second central moment of the operators SC_m defined by (5) is given by*

$$\left(SC_m\right)\left(\left(t-x\right)^2;x\right) = \frac{x}{b_m}\left(1-\frac{x}{b_m}\right)g_m\left(u\right)$$

with $u = 4\left(\frac{x}{b_m}\left(1-\frac{x}{b_m}\right)\right) \in [0,1]$ and $g_m\left(u\right)$ is defined by equation (6).

Proof. Starting with the equality

$$\sum_{k=0}^{m} p_{m,k}\left(\frac{x}{b_m}\right)e^{u\left(\frac{x}{b_m}-\frac{k}{m}\right)}e^{ik\theta} = e^{\frac{ux}{b_m}}\left(\frac{x}{b_m}e^{-\frac{u}{m}}e^{i\theta}+1-\frac{x}{b_m}\right)^m \tag{9}$$

and differentiating both sides of (9) with respect to u, and taking $u=0$, we obtain

$$\sum_{k=0}^{m} p_{m,k}\left(\frac{x}{b_m}\right)\left(\frac{x}{b_m}-\frac{k}{m}\right)e^{ik\theta} = \frac{x}{b_m}\left(\frac{x}{b_m}e^{i\theta}+1-\frac{x}{b_m}\right)^{m-1}\left(\frac{x}{b_m}e^{i\theta}-e^{i\theta}+1-\frac{x}{b_m}\right).$$

Using Parseval's equality, we have

$$\sum_{k=0}^{m} p_{m,k}^2\left(\frac{x}{b_m}\right)\left(\frac{x}{b_m}-\frac{k}{m}\right)^2 = \frac{8\left(\frac{x}{b_m}\left(1-\frac{x}{b_m}\right)\right)^2}{\pi}\int_0^{\pi/2}\left(1-4\frac{x}{b_m}\left(1-\frac{x}{b_m}\right)\sin^2\theta\right)^{m-1}\sin^2\theta d\theta. \tag{10}$$

Here, taking $\sin^2\theta = t$ in (10), we obtain

$$\sum_{k=0}^{m} p_{m,k}^2\left(\frac{x}{b_m}\right)\left(\frac{x}{b_m}-\frac{k}{m}\right)^2 = \frac{4\left(\frac{x}{b_m}\left(1-\frac{x}{b_m}\right)\right)^2}{\pi}\int_0^1 \frac{\left(1-4\frac{x}{b_m}\left(1-\frac{x}{b_m}\right)t\right)^{m-1}}{\sqrt{t\left(1-t\right)}}tdt. \tag{11}$$

With $u = 4\left(\frac{x}{b_m}\left(1-\frac{x}{b_m}\right)\right) \in [0,1]$, and from (5), (11) with (8), we have

$$\left(SC_m\right)\left(\left(t-x\right)^2;x\right) = \frac{x}{b_m}\left(1-\frac{x}{b_m}\right)g_m\left(u\right) \tag{12}$$

which completes the proof. □

3. Main results

Theorem 3.1. *The following inequality*

$$\frac{1}{2b_m^2}\frac{x\left(b_m-x\right)}{m} \leq \left(SC_m\right)\left(\left(t-x\right)^2;x\right), \tag{13}$$

is valid for all $x \in [0, b_m]$ and $m \geq 1$.

Proof. Since $\frac{1}{\sqrt{t(1-t)}} = \left(2\arcsin\sqrt{t}\right)'$, using integration by parts, we obtain

$$\begin{aligned}
\int_0^1 \frac{\left(1-ut\right)^m}{\sqrt{t\left(1-t\right)}}dt &= 2\int_0^1 \left(1-ut\right)^m\left(\arcsin\sqrt{t}\right)'dt \\
&= \left(1-u\right)^m\pi + 2mu\int_0^1\left(1-ut\right)^{m-1}\arcsin\sqrt{t}dt \\
&= \left(\varepsilon_m\left(u\right)+1\right)\ 2mu\int_0^1\left(1-ut\right)^{m-1}\arcsin\sqrt{t}dt \qquad (14)
\end{aligned}$$

where

$$\begin{aligned} \varepsilon_m(u) &= \frac{(1-u)^m \pi}{2mu\int\limits_0^1 (1-ut)^{m-1}\arcsin\sqrt{t}dt} \\ &< \frac{(1-u)^m \pi}{2mu\int\limits_0^1 (1-u)^{m-1}\arcsin\sqrt{t}dt} \\ &= \frac{2(1-u)}{mu}, \end{aligned} \tag{15}$$

for $u \in [0,1]$ and $m \geq 1$. From (14), using the inequality

$$\arcsin\sqrt{t} \leq \frac{t}{\sqrt{t(1-t)}}\ ,\ t \in [0,1)\,,$$

we obtain

$$\int\limits_0^1 \frac{(1-ut)^m}{\sqrt{t(1-t)}}dt \leq (\varepsilon_m(u)+1)\ \ 2mu\int\limits_0^1 \frac{t(1-ut)^{m-1}}{\sqrt{t(1-t)}}dt. \tag{16}$$

From (16) and (15), we deduce that:

$$mg_m(u) \geq \frac{1}{2(\varepsilon_m(u)+1)} \geq \frac{1}{2\left(\frac{2(1-u)}{mu}+1\right)}$$

hence

$$mg_m(u) \geq \frac{1}{2} - \frac{1-u}{2+(n-2)u} = \frac{1}{2} - \frac{1-4\left(1-\frac{x}{b_m}\right)\frac{x}{b_m}}{2\left(1+2(n-2)\left(1-\frac{x}{b_m}\right)\frac{x}{b_m}\right)},$$

with $u := 4\left(\frac{x}{b_m}\left(1-\frac{x}{b_m}\right)\right) \in [0,1]$ and $m = 1,2,....$ Finally, in virtue of (12), we obtain:

$$\begin{aligned} m(SC_m)\left((t-x)^2;x\right) &= \frac{x}{b_m}\left(1-\frac{x}{b_m}\right)mg_m(u) \\ &\geq \frac{\frac{x}{b_m}\left(1-\frac{x}{b_m}\right)}{2} - \frac{\frac{x}{b_m}\left(1-\frac{x}{b_m}\right)\left(1-4\left(1-\frac{x}{b_m}\right)\frac{x}{b_m}\right)}{2\left(1+2(n-2)\left(1-\frac{x}{b_m}\right)\frac{x}{b_m}\right)} \end{aligned}$$

and the proof is complete. □

Corollary 3.2. *The second central moments of the operators SC_m defined by (5) satisfy the following inequalities:*

$$\frac{1}{2b_m^2}\frac{x(b_m-x)}{m} \leq (SC_m)\left((t-x)^2;x\right) \leq p\ M_p\ \frac{x(b_m-x)}{m}, \tag{17}$$

$$(SC_m)\left((t-x)^2;x\right) \leq p\ M_p\ \ (C_m)\left((t-x)^2;x\right)$$

for $x \in [0,b_m]$ and $1 \leq p \leq m$.

Proof. Using (12) and (7), we have

$$\begin{aligned} (SC_m)\left((t-x)^2;x\right) &= \frac{x}{b_m}\left(1-\frac{x}{b_m}\right)g_m(u) \\ &\leq M_m\frac{x}{b_m}\left(1-\frac{x}{b_m}\right) \\ &\leq p\ M_p\ \frac{x(b_m-x)}{m} \\ &\leq p\ M_p\ \ (C_m)\left((t-x)^2;x\right) \end{aligned} \tag{18}$$

for all $1 \leq p \leq m$. Putting together (13) and (18), we get the desired result which states that the second central moments $(SC_m)\left((t-x)^2;x\right)$ of the operators SC_m are of order exactly

$\frac{x(b_m-x)}{m}$, and are smaller than the corresponding ones of the classical Bernstein-Chlodovsky operators. $\square$

In the approximation of a function by positive linear operators not only the convergence of operators is required but also the speed of convergence is important. The rate of convergence depends on the smoothness properties of the function and appropriate tool for estimating the smoothness of function are represented by modulus of continuity. We compute the rate of convergence of the constructed operators in terms of modulus of continuity and class of Lipschitz function.

Theorem 3.3. *For any $f \in C_B[0,\infty)$ (space of all bounded and continuous functions on $[0,\infty)$), we have the estimate*

$$\left|(SC_m)(f;x) - f(x)\right| \le \left(1 + \sqrt{p\ M_p}\right)\omega\left(f; \sqrt{\frac{x(b_m-x)}{m}}\right)$$

where $1 \le p \le m$, M_m is given by (7) and $\omega(f;.)$ is the modulus of continuity.

Proof. Since

$$\begin{aligned} \left|(SC_m)(f;x) - f(x)\right| &= \left|\sum_{k=0}^{m} f\left(\frac{b_m}{m}k\right) c_{m,k}\left(\frac{x}{b_m}\right) - f(x)\right| \\ &\le \sum_{k=0}^{m} c_{m,k}\left(\frac{x}{b_m}\right)\left|f\left(\frac{b_m}{m}k\right) - f(x)\right|, \end{aligned}$$

by using Cauchy-Schwartz inequality and the property

$$\left|f(t) - f(x)\right| \le \omega(f;\delta)\left(1 + \frac{|t-x|}{\delta}\right),$$

we get

$$\begin{aligned} \left|(SC_m)(f;x) - f(x)\right| &\le \sum_{k=0}^{m} c_{m,k}\left(\frac{x}{b_m}\right)\left\{1 + \frac{1}{\delta}\left|\frac{b_m}{m}k - x\right|\right\}\omega(f;\delta) \\ &\le \left\{1 + \frac{1}{\delta}\left(\sum_{k=0}^{m} c_{m,k}\left(\frac{x}{b_m}\right)\left(\frac{b_m}{m}k - x\right)^2\right)^{1/2}\left(\sum_{k=0}^{m} c_{m,k}\left(\frac{x}{b_m}\right)\right)^{1/2}\right\}\omega(f;\delta) \\ &\le \left\{1 + \frac{1}{\delta}\left((SC_m)\left((.-x)^2;x\right)\right)^{1/2}\left((SC_m)(1;x)\right)^{1/2}\right\}\omega(f;\delta). \end{aligned}$$

Using the fact that $(SC_m)(1;x) = 1$ and (17), we have

$$\left|(SC_m)(f;x) - f(x)\right| \le \left\{1 + \frac{1}{\delta}\left(p\ M_p\ \frac{x(b_m-x)}{m}\right)^{1/2}\right\}\omega(f;\delta),$$

choosing $\delta = \sqrt{\frac{x(b_m-x)}{m}}$, we obtain the result. $\square$

Corollary 3.4. *For any $f \in C_B[0,\infty)$, the sequence of operators $(SC_m f)_{m\in\mathbb{N}}$ converges uniformly to f on $[0,\infty)$.*

Theorem 3.5. *Let $f \in C_B[0,\infty)$, $M > 0$ and $0 < \mu \le 1$, then*

$$\left|(SC_m)(f;x) - f(x)\right| \le M\left(\delta_m(x)\right)^{\frac{\mu}{2}}$$

for each $f \in Lip_M(\mu) = \{f : |f(\eta_1) - f(\eta_2)| \le M|\eta_1 - \eta_2|^\mu, \eta_1, \eta_2 \in [0,\infty)\}$ and for $\delta_m(x) = (SC_m)\left((t-x)^2;x\right)$

Proof. We prove this theorem by using the definition of Lipschitz function and Hölder's inequality.

$$\begin{aligned} \left|(SC_m)(f;x) - f(x)\right| &\le SC_m\left(|f(t) - f(x)|;x\right) \\ &\le M\ SC_m\left(|t-x|^\mu;x\right). \end{aligned}$$

Therefore

$$\begin{aligned}
\left|(SC_m)(f;x)-f(x)\right| &\leq M\sum_{k=0}^{m}c_{m,k}\left(\frac{x}{b_m}\right)\left|\frac{b_m}{m}k-x\right|^{\mu}\\
&\leq M\sum_{k=0}^{m}\left(c_{m,k}\left(\frac{x}{b_m}\right)\right)^{\frac{2-\mu}{2}}\left(c_{m,k}\left(\frac{x}{b_m}\right)\right)^{\frac{\mu}{2}}\left|\frac{b_m}{m}k-x\right|^{\mu}\\
&\leq M\left\{\left(\sum_{k=0}^{m}c_{m,k}\left(\frac{x}{b_m}\right)\right)^{\frac{2-\mu}{2}}\left(\sum_{k=0}^{m}c_{m,k}\left(\frac{x}{b_m}\right)\left|\frac{b_m}{m}k-x\right|^{2}\right)^{\frac{\mu}{2}}\right\}\\
&= M\left((SC_m)\left((.-x)^2;x\right)\right)^{\frac{\mu}{2}}.
\end{aligned}$$

Choosing $\delta_m\left(\,x\right)=(SC_m)\left((.-x)^2;x\right)$, proof is completed. □

References

[1] Abel U. and Kushnirevych V., (2019), Voronovskaja type theorems for positive linear operators related to squared Bernstein polynomials, Positivity, 23, 697-710.

[2] Abel U., (2019/20), Voronovskaja type theorems for positive linear operators related to squared fundamental functions, in: Constructive Theory of Functions, Sozopol, (B. Draganov, K. Ivanov, G. Nikolov and R. Uluchev, Eds.), 1-21, Prof. Marin Drinov Publishing House of BAS, Sofia.

[3] Bernstein S.N., (1912/13), Demonstration du Thěoreme de Weierstrass fonděe sur le calcul des probabilitěs, Comm. Soc. Math. Kharkow 13,1-2.

[4] Chlodovsky I., (1937), Sur le děveloppement des functions děfinies dans un intervalle infini en sěries de polynomes de M. S. Bernstein, Compositio Math., 4, 380-393.

[5] Gavrea I. and Ivan M., (2014), On a conjecture concerning the sum of the squared Bernstein polynomials, Applied Mathhematics and Computation, 241, 70-74.

[6] Gavrea I. and Ivan M., (2017), On a new sequence of positive linear operators related to squared Bernstein polynomials, Positivity, 21, 911-917.

[7] Holhoş A., (2019), A sequence of positive linear operators related to powered Baskakov basis, Carpathian J. Math., 35(1), 51-58.

[8] Holhoş A., (2019), Voronovskaya theorem for a sequence of positive linear operators related to squared Bernstein polynomials, Positivity, 23, 571-580.

[9] Karsli H., (2011), Recent results on Chlodovsky operators, Stud. Univ. Babeş-Bolyai Math., 56(2), 423-436.

[10] Mohammad I.J. and Mohammad A.J., (2022), Neural Network of Multivariate Square Rational Bernstein Operators, Communications in Mathematics and Applications, 13(2), 585-594.

[11] Rahman S., Aral A. and Mursaleen M., (2023), Approximation of operators related to squared Szăsz-Mirakjan basis functions, Filomat, 37(7), 2141-2150.

Hüseyin Erhan Altin[1*]

Bolu Abant İzzet Baysal University, Faculty of Science and Arts, Department of Mathematics, 14280 Golkoy/Bolu

Harun Karsli[2]

Bolu Abant İzzet Baysal University, Faculty of Science and Arts, Department of Mathematics, 14280 Golkoy/Bolu

Email(s): erhanaltin@ibu.edu.tr[1*], karsli_h@ibu.edu.tr[2]

On new means generated by inverse of eigenfunctions of (p,q)-Laplacian

Barkat Ali Bhayo*[1] and József Sándor[2]

ABSTRACT. In this paper, we generalize Neuman-Sándor mean M and Seiffert mean P as an application of the inverse of eigenfunctions generalized hyperbolic functions with two parameters. Moreover, two-sided inequalities involving these generalized means are established.

2020 MATHEMATICS SUBJECT CLASSIFICATIONS: 34L10, 33E05, 33C75.

KEYWORDS: Generalized trigonometric functions, hypergeometric functions, p-Laplacian, logarithmic mean, two Seiffert means.

1. INTRODUCTION

In 1879, Lindberg [17] originally introduced generalized trigonometric functions. In 1995, Lindqvist [18] also introduced these functions, his work was highly cited by several authors. The recent literature on these functions includes several dozens of papers. These functions were used to study problems of existence, bifurcation and oscillation of solutions of differential equations, applications to differential equations involving the p-Laplacian, simple generalization of the classical trigonometric and hyperbolic functions, generalization of elliptic integrals of the first and the second kind, and generalization of means of two variables. The reader is referred to see [5, 7, 6, 11, 14, 27, 29, 31] and the bibliography therein.

The *Gaussian hypergeometric function* is defined by

$$F\left(a,b;c;z\right) = {}_2F_1\left(a,b;c;z\right) = \sum_{n=0}^{\infty} \frac{(a,n)(b,n)}{(c,n)} \frac{z^n}{n!}, \quad |z|<1,$$

where (a,n) denotes the shifted factorial function

$$(a,n) = a(a+1)(a+2)\dots(a+n-1), \quad n=1,2,3,\dots,$$

and $(a,0)=1$ for $a \neq 1$. The reader is referred to see [4] for the applications of Gaussian hypergeometric function in various fields of the mathematical and natural sciences.

Special functions, such as classical *gamma function* Γ, the *digamma function* ψ and the *beta function* $B(.,.)$ have close relation with hypergeometric function. For $x,y>0$, these functions are defined by

$$\Gamma(x) = \int_0^\infty e^{-t}t^{x-1}\,dt, \quad \psi(x) = \frac{\Gamma'(x)}{\Gamma(x)}, \quad B(x,y) = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)},$$

respectively. The hypergeometric function can be represented in the integral form as follows

$$F(a,b;c;z) = \frac{\Gamma(c)}{\Gamma(b)(c-b)} \int_0^1 t^{b-1}(1-t)^{c-b-1}(1-zt)^{-a}dt. \tag{1}$$

The eigenfunction $\sin_p$ of the so-called one-dimensional p-Laplacian problem [15]

$$-\Delta_p u = -\left(|u'|^{p-2}u'\right)' = \lambda|u|^{p-2}u,\ u(0)=u(1)=0, \quad p>1,$$

is the inverse function of $F_p : [0,1] \to \left[0, \frac{\pi_p}{2}\right]$, defined as

$$F_p(x) = \operatorname{arcsin}_p(x) = \int_0^x (1-t^p)^{-\frac{1}{p}}\,dt,$$

where

$$\pi_p = 2\operatorname{arcsin}_p(1) = \frac{2}{p}\int_0^1 (1-s)^{-\frac{1}{p}} s^{\frac{1}{p}-1} ds = \frac{2}{p}\, B\left(1-\frac{1}{p}, \frac{1}{p}\right) = \frac{2\pi}{p\,\sin\left(\frac{\pi}{p}\right)}.$$

The function $\arcsin_p$ is called the generalized inverse sine function, and its inverse function $\sin_p : [0, \pi_p/2] \to [0,1]$ is called generalized sine function. For $x \in [\pi_p/2, \pi_p]$, one can extends the function $\sin_p$ to $[0, \pi_p]$ by defining $\sin_p(x) = \sin(\pi_p - x)$, and further extension can be achieved on $\mathbb{R}$ by oddness and 2π-periodicity. The range of p is restricted to $(1, \infty)$ because only in this case $\sin_p(x)$ can be made periodic like usual sine function.

Similarly, the other generalized inverse trigonometric and hyperbolic functions $\arccos_p : (-1,1) \to (-a_p, a_p)$, $\arctan_p : (-\infty, \infty) \to (-a_p, a_p)$, $\mathrm{arcsinh}_p : (-\infty, \infty) \to (-\infty, \infty)$, $\mathrm{arctanh}_p : (-1,1) \to (-\infty, \infty)$ are defined as follows

$$
\begin{aligned}
\arccos_p(x) &= \int_0^{(1-x^p)^{\frac{1}{p}}} (1-|t|^p)^{-\frac{1}{p}} dt, \quad \arctan_p(x) = \int_0^x (1+|t|^p)^{-1} dt, \\
\mathrm{arcsinh}_p(x) &= \int_0^x (1+|t|^p)^{-\frac{1}{p}} dt, \quad \mathrm{arctanh}_p(x) = \int_0^x (1-|t|^p)^{-1} dt,
\end{aligned} \tag{2}
$$

where $a_p = \pi_p/2$. Above inverse generalized trigonometric and hyperbolic functions coincide with usual trigonometric and hyperbolic functions for $p = 2$.

Generalization of means with two parameters: For two positive real numbers a and b, we define arithmetic mean A, geometric mean G, logarithmic mean L, two Seiffert means P and T, and Neuman-Sándor mean M introduced in [21] as follows,

$$A = A(a,b) = \frac{a+b}{2}, \quad G = G(a,b) = \sqrt{ab},$$

$$L = L(a,b) = \frac{a-b}{\log(a) - \log(b)}, \quad a \neq b,$$

$$P = P(a,b) = \frac{a-b}{2\mathrm{arcsin}\left(\frac{a-b}{a+b}\right)},$$

$$T = T(a,b) = \frac{a-b}{2\mathrm{arctan}\left(\frac{a-b}{a+b}\right)},$$

$$M = M(a,b) = \frac{a-b}{2\mathrm{arcsinh}\left(\frac{a-b}{a+b}\right)}.$$

Neuman [23, 24] generalized the logarithmic mean L, two Seiffert means P and T, and the Neuman-Sándor mean M by introducing the the p-version of the Schwab-Borchardt mean SB_p as follows

$$L_p = L_p(a,b) = SB_p(A_{p/2}, G) = \frac{A_{p/2} v_p}{\mathrm{arctanh}_p(v_p)},$$

$$P_p = P_p(a,b) = SB_p(G, A_{p/2}) = \frac{A_{p/2} v_p}{\mathrm{arcsin}_p(v_p)},$$

$$T_p = T_p(a,b) = SB_p(A_{p/2}, A_p) = \frac{A_{p/2} v_p}{\mathrm{arctan}_p(v_p)},$$

$$M_p = M_p(a,b) = SB_p(A_p, A_{p/2}) = \frac{A_{p/2} v_p}{\mathrm{arcsinh}_p(v_p)},$$

where

$$SB_p(a,b) = b\, F\left(\frac{1}{p}, \frac{1}{p}; 1 + \frac{1}{p}, 1 - \left(\frac{a}{b}\right)^p\right)^{-1},$$

$$v_p = \frac{|x^{p/2} - y^{p/2}|}{x^{p/2} + y^{p/2}},$$

and $A_p = A_p(a,b)$ is a power mean of order p.

Motivated by the work of Neuman, Bhayo and Sándor gave a natural and new generalization of L, P, T and M in [10, Theorem 2.1] by utilizing the generalized trigonometric and generalized hyperbolic functions as follows:

For $p \geq 2$ and $a > b > 0$, the following functions define a mean of two variables a and b

$$(3)\quad \begin{aligned}
\tilde{P}_p = \tilde{P}_p(a,b) &= \frac{a-b}{2\mathrm{arcsin}_p\left(\frac{a-b}{a+b}\right)} = \frac{x}{\mathrm{arcsin}_p(x)}A = \frac{A}{F\left(\frac{1}{p},\frac{1}{p};1+\frac{1}{p};x^p\right)},\\
\tilde{T}_p = \tilde{T}_p(a,b) &= \frac{a-b}{2\mathrm{arctan}_p\left(\frac{a-b}{a+b}\right)} = \frac{x}{\mathrm{arctan}_p(x)}A = \frac{A\cdot(1+x^p)^{1/p}}{F\left(\frac{1}{p},\frac{1}{p};1+\frac{1}{p};\frac{x^p}{1+x^p}\right)},\\
\tilde{L}_p = \tilde{L}_p(a,b) &= \frac{a-b}{2\mathrm{artanh}_p\left(\frac{a-b}{a+b}\right)} = \frac{x}{\mathrm{artanh}_p(x)}A = \frac{A}{F\left(1,\frac{1}{p};1+\frac{1}{p};x^p\right)},\\
\tilde{M}_p = \tilde{M}_p(a,b) &= \frac{a-b}{2\mathrm{arsinh}_p\left(\frac{a-b}{a+b}\right)} = \frac{x}{\mathrm{arsinh}_p(x)}A = \frac{A\cdot(1+x^p)^{1/p}}{F\left(1,\frac{1}{p};1+\frac{1}{p};\frac{x^p}{1+x^p}\right)},
\end{aligned}$$

where $x = (a-b)/(a+b)$.

In [27], generalized inverse trigonometric function arcsin_p was extended for two parameters $p, q > 1$ as follows,

$$\mathrm{arcsin}_{p,q}(x) = \int_0^x (1-t^q)^{-\frac{1}{p}}dt,$$

for $x \in (0,1)$.

Letting $t = z^{1/q}$, we observe that

$$\mathrm{arcsin}_{p,q}(x) = \frac{1}{q}\int_0^{x^q} z^{1/q-1}(1-z)^{-1/p}dz = \frac{1}{q}\tilde{B}\left(\frac{1}{q},\, 1-\frac{1}{p},\, x^q\right),$$

where $\tilde{B}(a,b,x)$ is incomplete beta function defined as

$$\tilde{B}(a,b,x) = \int_0^x t^{a-1}(1-t)^{b-1}dt.$$

The function $\mathrm{arcsin}_{p,q}(x)$ is the inverse function of $\sin_{p,q}$, defined on the the interval $[0, \pi_{p,q}/2]$, where

$$\pi_{p,q} = 2\mathrm{arcsin}_{p,q}(1) = \frac{2}{q}\,B\left(1-\frac{1}{p},\frac{1}{q}\right) = \frac{2\pi}{q\,\sin\left(\frac{\pi}{p}\right)}.$$

For $T = \pi_{p,q}$ the function $u(t) = \sin_{p,q}(t)$ is a solution to the following problem considered by Drábek and Manásevich Let $\phi_p(x) = |x|^{p-2}x$. For $T, \lambda > 0$ and $p, q > 1$

$$\begin{cases} (\phi_p(u'))' + \lambda\,\phi_q(u) = 0, & t \in (0,T),\\ u(0) = u(T) = 0. \end{cases}$$

For $x \in (0,\,1)$, we also define $\mathrm{arccos}_{p,q}(x) = \mathrm{arcsin}_{p,q}((1-x^p)^{1/q})$, and

$$\cos_{p,q}(x) = \frac{d}{dx}\sin_{p,q}(x),\; x \in [0,\, \pi_{p,q}/2].$$

Letting $y = \sin_{p,q}(x)$, we get $\cos_{p,q}(x) = (1-(\sin(x)^q))^{1/p}$ and

$$|\cos_{p,q}|^p + |\sin_{p,q}|^q = 1\,.$$

Similarly, the generalized tangent function is defined as $\tan_{p,q} = (\sin_{p,q}(x))/(\cos_{p,q}(x))$, and its inverse is denoted by $\mathrm{arctan}_{p,q}$, see [9].

Motivated by the work of Takeuchi [27], Bhayo and Vuorinen [11] defined the generalized inverse hyperbolic sine function $\mathrm{arcsinh}_{p,q}$ as follows

$$\mathrm{arcsinh}_{p,q}(x) = \int_0^x (1+t^q)^{-\frac{1}{p}}dt = x\,F\left(\frac{1}{p},\frac{1}{q};1+\frac{1}{q};x^q\right),$$

for $x \in (0,1)$. In similar fashion, generalized hyperbolic cosine and tangent function are defined by

$$\cosh_{p,q}(x) = \frac{d}{dx}\sinh_{p,q}(x), \quad \tanh_{p,q}(x) = \frac{\sinh_{p,q}(x)}{\cosh_{p,q}(x)}, \quad x \geq 0,$$

and the inverse of $\tanh_{p,q}$ is denoted by arctanh. If follows from the definition that

$$|\cosh_{p,q}|^p - |\sinh_{p,q}|^q = 1\,.$$

In [11], it was proved that for $p, q > 1$ and $x \in (0, 1)$ one has

$$x\left(1+\frac{x^q}{p(1+q)}\right) < \operatorname{arcsin}_{p,q}(x) < \min\{\frac{\pi_{p,q}}{2}, (1-x^q)^{-1/(p(1+q))}x\} \tag{4}$$

$$\left(\frac{x^p}{1+x^q}\right)^{1/p} r_{p,q}(x) < \operatorname{arcsinh}_{p,q}(x) < \left(\frac{x^p}{1+x^q}\right)^{1/p} s_{p,q}(x), \tag{5}$$

where

$$r_{p,q}(x) = \max\left\{\left(1-\frac{qx^q}{p(1+q)(1+x^q)}\right)^{-1}, (1+x^q)^{1/p}\left(\frac{pq+p+qx^q}{p(1+q)}\right)^{-1/q}\right\}$$

and $s_{p,q}(x) = (1 - x^p/(1 + x^q))^{-q/(p(1+q))}$.

It is easy to observe that (p, q)-functions reduce to p-functions for $p = q$, and $\pi_{p,q}$ to π_p. It follows from definition that $\pi_{p,q} \le \pi_p$ for $q \ge p > 1$.

1.1. **Preliminaries.**

Lemma 1.1. [3, Theorem 2] *For $-\infty < a < b < \infty$, let $f, g : [a, b] \to \mathbb{R}$ be continuous on $[a, b]$, and be differentiable on (a, b). Let $g'(x) \neq 0$ on (a, b). If $f'(x)/g'(x)$ is increasing (decreasing) on (a, b), then so are*

$$\frac{f(x)-f(a)}{g(x)-g(a)} \quad \text{and} \quad \frac{f(x)-f(b)}{g(x)-g(b)}.$$

If $f'(x)/g'(x)$ is strictly monotone, then the monotonicity in the conclusion is also strict.

Lemma 1.2. [9] *For $p, q > 1$, the following inequalities hold true,*

(1) $\dfrac{x}{\operatorname{arcsin}_{p,q}(x)} < \dfrac{\sin_{p,q}(x)}{x}, \quad x \in (0, 1),$

(2) $\dfrac{x}{\operatorname{arcsinh}_{p,q}(x)} < \dfrac{\sinh_{p,q}(x)}{x}, \quad x \in (0, \infty),$

(3) $\dfrac{x}{\operatorname{arctan}_{p,q}(x)} < \dfrac{\tan_{p,q}(x)}{x}, \quad x \in (0, 1),$

(4) $\dfrac{x}{\operatorname{arctanh}_{p,q}(x)} < \dfrac{\tanh_{p,q}(x)}{x}, \quad x \in (0, 1).$

For easy reference we recall some well-known inequalities from the literature as follows.

Cauchy-Bouniakowski inequality. If $f, g : [a, b] \to \mathbb{R}$ are integrable, then

$$\left(\int_a^b f(x)g(x)dx\right)^2 \le \int_a^b f(x)^2dx \int_a^b g(x)^2dx. \tag{6}$$

Pólya-Szegő inequality. If $f, g : [a, b] \to \mathbb{R}$ are integrable, and for all $x \in [a, b]$

$$0 < \alpha < f(x) < A, \quad 0 < \beta < g(x) < B,$$

then

$$\frac{\int_a^b f(x)^2dx \int_a^b g(x)^2dx}{\left(\int_a^b f(x)g(x)dx\right)^2} \le K(\alpha, A, \beta, B), \tag{7}$$

where

$$K = K(\alpha, A, \beta, B) = \frac{1}{4}\left(\sqrt{\frac{AB}{\alpha\beta}} + \sqrt{\frac{\alpha\beta}{AB}}\right)^2.$$

Minkowski's inequality. Let $f, g : [a, b] \to \mathbb{R}$ be integrable and $f, g > 0$. Write

$$h_t(f) = \left(\int_a^b f(x)^t dx\right)^{1/t}.$$

Then one has

$$h_t(f+g) \le h_t(f) + h_t(g), \quad \text{for} \quad t \ge 1, \tag{8}$$

$$h_t(f+g) \ge h_t(f) + h_t(g), \quad \text{for} \quad t \le 1. \tag{9}$$

Diaz-Metcalf inequality. Let $f, g : [a, b] \to \mathbb{R}$ be integrable and suppose that there exist constants m and M such that

$$m \le g(x)/f(x) \le M.$$

Then one has

$$\int_a^b g^2(x)dx + m \cdot M \cdot \int_a^b f^2(x)dx \le (m + M) \cdot \int_a^b f(x)g(x)dx. \tag{10}$$

The paper is organized as follows. In section 1, we give the definition of the special functions involved in our formulation of the main results, and definition of generalized means with one parameter and the statement their previous results. Few lemmas and some well-known inequalities are given in the subsection (‘preliminaries). In section 2, we give the definition of the generalized means with two parameters and the statement of the main results in the form of theorems. Section 3 is consisting of the proof of theorems. In Section 4, we present a conjecture.

2. Main results

The mean $\tilde{P}_p$ and $\tilde{M}_p$ can be further generalized for two parameters $p, q \ge 2$ as follows,

$$\begin{aligned}
\tilde{P}_{p,q} = \tilde{P}_{p,q}(a,b) &= \frac{a-b}{2\mathrm{arcsin}_p\left(\frac{a-b}{a+b}\right)} = \frac{x}{\mathrm{arcsin}_{p,q}(x)}A, \\
\tilde{M}_{p,q} = \tilde{M}_{p,q}(a,b) &= \frac{a-b}{2\mathrm{arsinh}_p\left(\frac{a-b}{a+b}\right)} = \frac{x}{\mathrm{arsinh}_{p,q}(x)}A \\
\tilde{T}_{p,q} = \tilde{T}_p(a,b) &= \frac{a-b}{2\mathrm{arctan}_{p,q}\left(\frac{a-b}{a+b}\right)} = \frac{x}{\mathrm{arctan}_{p,q}(x)}A \\
\tilde{L}_{p,q} = \tilde{L}_p(a,b) &= \frac{a-b}{2\mathrm{artanh}_{p,q}\left(\frac{a-b}{a+b}\right)} = \frac{x}{\mathrm{artanh}_{p,q}(x)}A,
\end{aligned} \tag{11}$$

where $x = (a-b)/(a+b)$ with $0 < b < a$. Clearly, for $p = q$ we have $\tilde{P}_{p,p} = \tilde{P}_p$, $\tilde{M}_{p,p} = M_p$, $\tilde{T}_{p,p} = T_p$, and $\tilde{L}_{p,p} = L_p$ For $p = 2$, $\tilde{P}_2 = P$, $\tilde{T}_2 = T$, $\tilde{L}_2 = L$, $\tilde{M}_2 = M$, $\tilde{T}_2 = T$, and $\tilde{L}_2 = M$.

Theorem 2.1. *For $p, q \ge 2$ and $a > b > 0$, the function $\tilde{P}_{p,q}$ and $\tilde{M}_{p,q}$ define a mean function of two variables a and b.*

Theorem 2.2. *For $2 \le p < q$ and $x = (a-b)/(a+b)$ with $0 < b < a$, we have*

$$P_{p,q}\, M_{p,q} \le (P_{2p,2q})^2 \le k(x,p,q) P_{p,q}\, M_{p,q}, \tag{12}$$

where

$$k(x,p,q) = \frac{((1+x^q)^{2/p} + (1-x^q)^{2/p})^2}{4((1-x^{2q})^{1/p}}.$$

Theorem 2.3. *For $a > b > 0$ and $x = (a-b)/(a+b)$, we have*

$$\frac{1}{\tilde{P}_{p,q}} + \frac{r}{\tilde{M}_{p,q}} \le \frac{r(p,q,x)+1}{\tilde{P}_{2p,2q}}, \tag{13}$$

$$(\tilde{P}_{2p,2q})^{2p}\left(\frac{1}{(\tilde{P}_{p,q})^p} + \frac{1}{(\tilde{M}_{p,q})^p}\right) \le R(p,q,x), \tag{14}$$

where

$$r(p,q,x) = \frac{(1+x^q)}{(1-x^q)}^{1/(2p)},$$

$$R(p,q,x) = [(1-x^{2q})^{1/(2p)} + (1-x^{2q})^{-1/(2p)}]^{1/(2p)}/2^{2p-1}.$$

3. Proof of main results

Proof of Theorem 2.1. Clearly, for $t \in (0,1)$ and $2 < p < q$, the following inequality
$$\int_0^x (1+t^q)^{-1/p}dt < \int_0^x (1-t^q)^{-1/p}dt$$
implies
$$\tilde{P}_{p,q} < \tilde{M}_{p,q}. \tag{15}$$
Similarly, for $t \in (0,1)$ and $2 < p < q$ the inequality
$$\int_0^x (1+t^q)^{-1/p}dt < \int_0^x (1+t^p)^{-1/p}dt$$
and
$$\int_0^x (1-t^{-q})^{-1/p}dt > \int_0^x (1-t^{-p})^{-1/p}dt$$
yield
$$\tilde{M}_{p,q} > \tilde{M}_p \quad \text{and} \quad \tilde{P}_{p,q} < \tilde{P}_p. \tag{16}$$
Since $\tilde{P}_{p,q}$ is a mean, by (15) to prove that $\tilde{M}_{p,q}$ is a mean, it is sufficient to find an upper bound for $\tilde{M}_{p,q}$, which is a mean.

Let $Q = Q(a,b) = \sqrt{(a^2+b^2)/2}$. Let $(a-b)/(a+b) = x$. Then it is easy to see that one has the following identity:
$$Q/A = \sqrt{1+x^2}. \tag{17}$$
Now, since
$$\frac{A}{\tilde{M}_{p,q}} = \frac{1}{x}\int_0^x (1+t^q)^{-1/p}dt.$$
Therefore, by (17), to prove that $A/Q <= A/\tilde{M}_{p,q}$, it is sufficient to prove that
$$(1+t^q)^{1/q} \le (1+x^2)^{1/2}. \tag{18}$$
For $q \ge 2$ and $t \in [0,1]$, we have
$$(1+t^q)^{1/p} \le (1+t^2)^{1/p} \le (1+x^2)^{1/p} \le (1+x^2)^{1/2},$$
so (18) follows. This implies that $A/\tilde{M}_{p,q} \ge A/Q$, so
$$\tilde{M}_p, q \le Q. \tag{19}$$
By (15) an (19) it follows that $\tilde{M}_{p,q}$ is also a mean. □

An other proof of $\tilde{P}_{p,q}$ mean. By definition, we have
$$\frac{\tilde{P}_{p,q}}{A} = \frac{2z}{\operatorname{arcsin}_{p,q}(z)},$$
where $z = (x-y)/(x+y)$ and $x > y$. Now, from the inequalities
$$x < \operatorname{arcsin}_{p,q}(t) < \frac{\pi_{p,q}}{2}x,$$
we get the double inequality:
$$\frac{2}{\pi_{p,q}}A < \tilde{P}_{p,q} < A, \tag{20}$$
assume that $\alpha_{p,q} \ge 2$. Now, by Jordan's inequality $\sin x > (2/\pi)x > x/2$, we get $\pi_{p,q} \le \pi_p < 4$. Let $0 < y < x$. Then from (20) we get
$$y < \tilde{P}_{p,q} < x,$$
where the right hand side is trivial, as $A = (x+y)/2 < x$, and for the left hand side, $(2/\pi_{p,q})A > (4/\pi_{p,q})y > y$, by $\pi_{p,q} < 4$. This implies that $\tilde{P}_{p,q}$ is a mean. □

The proof of the following corollary follows immediately from Lemma 1.2.

Corollary 3.1. *For $0 < y < x$, we have*

(1) $\tilde{P}_{p,q}(x,y) < 2\sin_{p,q}\left(\frac{x-y}{x+y}\right)\frac{A^2}{x-y}, \quad 2 \le p < q,$

(2) $\tilde{M}_{p,q}(x,y) < 2\sinh_{p,q}\left(\frac{x-y}{x+y}\right)\frac{A^2}{x-y}, \quad 2 \le p < q,$

Proof of Theorem 2.2. Let $f(t) = \sqrt{F(t)}$ and $g(t) = \sqrt{G(t)}$ in Cauchy-Bouniakowski inequality (6), where $F(t)$, $G(t) > 0$. Put $[a, b] = [0, x]$, then one gets the inequality:

$$\left(\int_0^x \sqrt{F(t)G(t)}dt\right)^2 \le \int_0^x F(t)dt \cdot \int_0^x G(t)dt. \tag{21}$$

With the same notations, from the Pólya-Szegö inequality (7) one gets:

$$k(x,p)\left(\int_0^x \sqrt{F(t)G(t)}dt\right)^2 \ge \int_0^x F(t)dt \cdot \int_0^x G(t)dt, \tag{22}$$

here $k(x, p)$ is as defined in Theorem 2.2. Let now $f(t) = (1-t^p)^{-1/p}$ and $g(t) = (1+t^p)^{-1/p}$. From (21) and (22) one obtains

$$\mathrm{arcsin}_{2p}(x)^2 \le \mathrm{arcsin}_p(x)\,\mathrm{arcsinh}_p(x), \tag{23}$$

and

$$\mathrm{arcsin}_p(x)\,\mathrm{arcsinh}_p(x) \le k(x,p)\mathrm{arcsin}_{2p}(x)^2, \tag{24}$$

respectively. By definition, inequality (23) and (24) imply the proof of left hand-side and right-hand side of (12), respectively. □

Proof of Theorem 2.3. Apply the Diaz-Metcalf inequality (10) for $f(t) = \sqrt{F(t)}$, $g(t) = \sqrt{G(t)}$, $[a, b] = [0, x]$, yielding

$$\int_0^x Gdt + M\cdot m\cdot \int_0^x Fdt \le (M+m)\cdot\int_0^x \sqrt{FG}dt.$$

Let $F(t) = (1+t^q)^{-1/p}$, $G(t) = (1-t^q)^{-1/p}$. Here $G(t)/F(t) = ((1+t^q)/(1-t^p))^{1/p}$, which is strictly increasing. Thus

$$m = 1 \le \sqrt{F/G} \le ((1+x^q)/(1-x^q))^{1/(2p)} = M.$$

One obtains

$$\mathrm{arcsin}_p(x) + M\cdot\mathrm{arcsinh}_{p,q}(x) \le (M+1)\mathrm{arcsin}_{2p,2q}(x), \tag{25}$$

this implies the proof of (13).

Let $[a, b] = [0, x]$ and $f(t) = (1+t^q)^{-1}$ and $g(t) = (1-t^q)^{-1}$. As $f(t)+g(t) = 2/(1-t^{2q})$, applying the Minkowski inequality (9) for $t = 1/p$, $p > 1$, we get

$$\mathrm{arcsin}_{p,q}(x)^p + \mathrm{arcsinh}_{p,q}(x)^p \le 2\left(\int_0^x A^2dt\right)^p, \tag{26}$$

where $A(t) = 1/(1-t^{2q})^{1/(2p)}$. Clearly, $\int_0^x A(t)dt = \mathrm{arcsin}_{2p,2q}(x)$, so for obtaining an upper bound for $\int_0^x A(t)^2dt$, we apply the Pólya-Szegő inequality for $f(t) = 1/(1-t^{2q})^{1/p}$ and $g(t) = 1$. Since in this case one has $1 \le f(t) \le 1/(1-x^{2q})^{1/p}$, we get from (7)

$$\int_0^x A(t)^2dt \le \mathrm{arcsin}_{2p,2q}(x)^2 R(x,p,q),$$

By using (26), finally we get

$$\frac{x^q(\mathrm{arcsin}_p(x)^p + \mathrm{arcsinh}_p(x)^p)}{\mathrm{arcsin}_{2p,2q}(x)^{2p}} \le R(p,q,x), \tag{27}$$

this implies inequality (14). □

Corollary 3.2. *For $2 \le p < q$, $x = (a-b)/(a+b)$ and $A = (a+b)/2$ with $0 < b < a$, we have*

$$\frac{x\,A}{\min\{\frac{\pi_{p,q}}{2}, (1-x^q)^{-1/(p(1+q))}x\}} < \tilde{P}_{p,q} < \frac{A}{\left(1+\frac{x^q}{p(1+q)}\right)} \tag{28}$$

$$\frac{(1+x^q)^{1/p}A}{s_{p,q}(x)} < \tilde{M}_{p,q} < \frac{(1+x^q)^{1/p}A}{r_{p,q}(x)}, \tag{29}$$

where $r_{p,q}(x)$ and $r_{p,q}(x)$ are as in (4).

Proof. Proof follows easily from (28) and (28). □

4. Conclusion

In this short paper we generalized the Seiffert mean P and Neuman-Sádor mean M with two parameters in the form of Theorem 2.1, and proved some inequalities involving these means. We finish the paper by posing the following conjecture.

Conjecture 4.1. *For $p, q \geq 2$ and $x = (a-b)/(a+b)$ with $0 < b < a$, the following functions*

$$\tilde{L}_{p,q} = \tilde{L}_{p,q}(a,b) = \frac{a-b}{2\operatorname{arctanh}_{p,q}\left(\frac{a-b}{a+b}\right)} = \frac{x}{\operatorname{arctanh}_{p,q}(x)}A,$$

$$\tilde{T}_{p,q} = \tilde{T}_{p,q}(a,b) = \frac{a-b}{2\operatorname{arctan}_{p,q}\left(\frac{a-b}{a+b}\right)} = \frac{x}{\operatorname{arctan}_{p,q}(x)}A,$$

define a mean functions of two variables a and b, where $\operatorname{arctanh}_{p,q}$ and $\operatorname{arctan}_{p,q}$ are the inverse functions of $\tanh_{p,q}$ and $\tan_{p,q}$, respectively.

References

[1] Abramowitz M., I. Stegun I., (1965), Handbook of Mathematical Functions with Formulas, Graphs and Mathematical Tables, National Bureau of Standards, Dover, New York.

[2] Anderson G.D., Vamanamurthy M.K. and Vuorinen M., (1997), Conformal Invariants, Inequalities and Quasiconformal Maps, J. Wiley.

[3] Anderson G.D., Vamanamurthy M.K. and Vuorinen M., (2006), Monotonicity Rules in Calculus, Amer. Math. Month. 113 (9) (2006) 805-816.

[4] Askey R., (1989), Handbooks of special functions, A Century of Mathematics in America, Part III (P. Duren, ed.), Amer. Math. Soc, Providence, RI, 369-391.

[5] Baricz Á., Bhayo B.A. and Pogány T.K., (2014), Functional inequalities for generalized inverse trigonometric and hyperbolic functions, J. Math. Anal. Appl., 417 (1), 2014 244-259.

[6] Baricz Á., Bhayo B.A. and Vuorinen M., (2015), Turán type inequalities for generalized inverse trigonometric functions, Filomat 29 (2) 303-313.

[7] Baricz Á., Bhayo B.A. and Klén K., (2015), Convexity properties of generalized trigonometric and hyperbolic functions, Aequat. Math. 89, 273-284.

[8] Bhayo B.A, Sándor J., (2014), On two new means of two variables II, Notes Number Th. Discr. Math. 20 (4) 1-10.

[9] Bhayo B.A, Sándor J., (2013), Inequalities connecting generalized trigonometric functions with their inverses, Issues of Analysis 2(20) 82-90.

[10] Bhayo B.A , Sándor J., (2019), On certain new means generated by generalized trigonometric functions, Tbilisi Math. J. 12(1) 1-16.

[11] Bhayo B.A., Vuorinen M., (2012), On generalized trigonometric functions with two parameters, J. Approx. Theory 164 (10) 1415-1426.

[12] Bhayo B.A., Vuorinen M., (2012), Inequalities for eigenfunctions of the p-Laplacian, Issues of Analysis, 2 (20) 13-35.

[13] Bushell P.J., Edmunds D.E, (2012), Remarks on generalised trigonometric functions, Rocky Mountain J. Math. 42 (1) 25-57.

[14] Edmunds E.D., Gurka P. and Lang L.,(2012), Properties of generalized trigonometric functions, J. Approx. Theory, 164 (1) 47-56.

[15] Drábek P., Manásevich R., (1999), On the closed solution to some $p-$Laplacian nonhomogeneous eigenvalue problems, Diff. and Int. Eqns. 12 723-740.

[16] Erdélyi A., Magnus W., Oberhettinger F. and Tricomi F.G., (1981), Higher Transcendental Functions, I, Melbourne.

[17] Lindqvist P., Peetre J., (2004), Comments on Erik Lundberg's 1879 thesis. Especially on the work of Gr̈an Dillner and his influence on Lundberg, Memorie dell'Instituto Lombardo (Classe di Scienze Matem. Nat.) 31.

[18] Lindqvist P., (1995), Some remarkable sine and cosine functions, Ricerche di Matematica, XLIV 269-290.

[19] Lindqvist P., Peetre J., (2003), p-arclength of the q-circle, The Mathematics Student 72 (1) 139-145.

[20] Mitrinović D.S., (1970), Analytic Inequalities, Springer-Verlag, Berlin.

[21] Neuman E., Sándor J.,(2003), On the Schwab-Borchardt mean, Math. Pannon., 14 (2) 253-266.

[22] Neuman E., Sándor J.,(2006), On the Schwab-Borchardt mean II, Math. Pannon. 17 (1) 49-59.

[23] Neuman E., (2014), On the p-version of the Schwab-Borchardt mean, Intern. J. Math. Math. Sci. 7 pp.

[24] Neuman E., (2015), On the p-version of the Schwab-Borchardt mean II, Intern. J. Math. Math. Sci., 4 pp.

[25] Sándor J., (2014), On two new means of two variables, Notes Number Th. Discr. Math. 20(1) 1-9.

[26] Sándor J., (2015), A note on certain bivariate means, Notes Number Th. Discr. Math. 20(5) 11-13.

[27] Takeuchi S., (2012), Generalized Jacobian elliptic functions and their application to bifurcation problems associated with p-Laplacian, J. Math. Anal. Appl. 385 (1) 24-35.

[28] Takeuchi S., (2019), Applications of generalized trigonometric functions with two parameters II, Diff. Eq. Appl., 4 (11) 563-575.

[29] Miyakawa H., Takeuchi S., (2021), Applications of a duality between generalized trigonometric and hyperbolic functions, J. Math. Anal. Appl. 502 (1), 25-41.
[30] Zhou S.-S., Qian W.-M., Chu Y.-M. and Zhang X.-H., (2015), Sharp power-type Heronian mean bounds for Sándor and Yang means, J. Inequal. Appl. 159 Doi:10.1186/s13660-015-0683-7.
[31] Yin L., Huang L-G., Wang Y-L. and Lin X-L., (2019), A survey for generalized trigonometric and hyperbolic functions, J. Math. Ineq. 13(3) 833-854.

Barkat Ali Bhayo (corresponding author)[1*]

LUT School of Energy Systems, Lappeenranta University of Technology, Mukkulankatu 19, 15210, Lahti, Finland

József Sándor[2]

Babeş-Bolyai University, Department of Mathematics, Str. Kogalniceanu nr. 1 400084 Cluj-Napoca, Romania

Email(s): barkat.bhayo@lut.fi[1*] (corresponding author), jsandor@math.ubbcluj.ro[2]

Stancu type Dunkl generalization of Szász-Durrmeyer operators involving two variable hermite polynomials

Zeynep Er*[1] , Harun Karslı[2] , Rabia Aktaş Karaman[3],
Fatma Taşdelen Yeşildal[4]

ABSTRACT. The aim of the present paper is to introduce a Stancu type Dunkl generalization of Szász- Durrmeyer operators involving two-variable Hermite polynomials defined by Krech [7]. Then, we give approximation properties for these operators with the help of Korovkin theorem. Furthermore, we obtain other approximation results via the class of Lipschitz functions, classical modulus of continuity, second ordermodulus of continuity and Peetre's K-functional.

2020 MATHEMATICS SUBJECT CLASSIFICATIONS: 41A25, 41A36, 33C45

KEYWORDS: Dunkl exponential, Szász operators, modulus of continuity, Lipschitz functions, two-variable Hermite polynomials, Szász- Durrmeyer operators.

1. INTRODUCTION

In 1912 , Bernstein [15] introduced the following sequence of operators $B_n : B[0,1] \to C[0,1]$ defined by

$$B_n(f,x) = \sum_{k=0}^{n} \binom{n}{k} x^k (1-x)^{n-k} f(\frac{k}{n}), \tag{1}$$

where $n \in N$, $x \in [0,1]$ and $f \in B[0,1]$. Some generalizations and modifications of the Bernstein polynomials can be found in [11], [9], [13]. Moreover, Szász [14] and Mirakjan [10] have introduced following operator

$$S_n(f,x) = e^{-nx} \sum_{k=0}^{\infty} \frac{(nx)^k}{k!} f(\frac{k}{n}), \tag{2}$$

where $n \in N$, $x \geq 0$, $f \in C[0,\infty)$. For $f \in C[0,\infty)$, Durrmeyer type integral modifications of the operators (2) was defined by Mazhar and Totik [9] as follows

$$D_n(f,x) = n \sum_{j=0}^{\infty} e^{-nx} \frac{(nx)^k}{k!} \int_0^{\infty} e^{-nx} \frac{(nx)^k}{k!} f(t)dt, \tag{3}$$

where $x \in [0,1]$. Also, Dunkl analogue of Szász operators given in [16] has been defined as

$$S_n^*(f,x) = \frac{1}{e_\mu(nx)} \sum_{k=0}^{\infty} \frac{(nx)^k}{\gamma_\mu(k)} f(\frac{k+2\mu\theta_k}{n}), \tag{4}$$

where $n \in N$, $x \geq 0$, $f \in C[0,\infty)$, $\mu \geq 0$.

The Dunkl exponential function is given for $\mu \geq -1/2$

$$e_\mu(nx) = \sum_{k=0}^{\infty} \frac{(nx)^k}{\gamma_\mu(k)}, \tag{5}$$

where the coefficients are as follows

$$\gamma_\mu(2k) = \frac{2^{2k}k!\Gamma(k+\mu+\frac{1}{2})}{\Gamma(\mu+\frac{1}{2})} \quad \text{and} \quad \gamma_\mu(2k+1) = \frac{2^{2k+1}k!\Gamma(k+\mu+\frac{3}{2})}{\Gamma(\mu+\frac{3}{2})} \tag{6}$$

and the recurrence relation

$$\frac{\gamma_\mu(k+1)}{\gamma_\mu(k)} = (2\mu\theta_{k+1}+k+1) \tag{7}$$

is satisfied where θ_k is given by [13]

$$\theta_k = \begin{cases} 0 & \text{if} \quad k=2p \\ 1 & \text{if} \quad k=2p+1 \end{cases} \tag{8}$$

for $p \in N \cup \{0\}$. The Dunkl generalization of two-variable Hermite polynomials [5] is defined by

$$\sum_{k=0}^{\infty} \frac{h_k^\mu(n,\alpha)}{\gamma_\mu(k)} x^k = e^{\alpha x^2} e_\mu(nx), \tag{9}$$

where

$$h_k^\mu(n,\alpha) = \frac{\gamma_\mu(k) H_k^\mu(n,a)}{k!}. \tag{10}$$

Also, Wafi and Rao [18] defined the Szász-Gamma operators based on Dunkl analogue as follows

$$D_n f(x) = \frac{1}{e_\mu(nx)} \sum_{k=0}^{\infty} \frac{(nx)^k}{\gamma_\mu(k)} \frac{n^{k+2\mu\theta_k+\lambda+1}}{\Gamma(k+2\mu\theta_k+\lambda+1)} \int_0^\infty t^{k+2\mu\theta_k+\lambda} e^{-nx} f(t)dt, \tag{11}$$

where $x \in [0,\infty)$, $\lambda \geq 0$ and the well-known Gamma function $\Gamma(x)$ is defined as

$$\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt. \tag{12}$$

Moreover, Dunkl-Gamma type operator in terms of generalization of two-variable Hermite polynomials [6] is defined by

$$S_n(f,x) = \frac{1}{e^{\alpha x^2} e_\mu(nx)} \sum_{k=0}^{\infty} \frac{h_k^\mu(n,\alpha)}{\gamma_\mu(k)} x^k \frac{n^{k+2\mu\theta_k+\lambda+1}}{\Gamma(k+2\mu\theta_k+\lambda+1)} \int_0^\infty t^{k+2\mu\theta_k+\lambda} e^{-nt} f(t)dt \tag{13}$$

where $\alpha \geq 0$, $\mu \geq -1/2$, $\lambda \geq 0$ and $x \in [0,\infty)$.

In the present paper, we first give some lemmas and definitions to obtain convergence properties of the operators. Then, we define the Stancu type Dunkl generalization of Szász-Durrmeyer operators involving two-variable Hermite polynomials. Finally, we give the rates of convergence of the operator using the classical modulus of continuity, second order modulus of continuity, Peetre's K-functional and in terms of the elements of the Lipschitz class $Lip_M(v)$.

2. Some Lemmas and Definitions

Lemma 2.1. *Let $h_k^\mu(n,\alpha)$ be the Dunkl generalization of two-variable Hermite polynomials. Then the following equalities are hold.*

$$i)\ \sum_{k=0}^{\infty}\frac{h^{\mu}_{k+1}(n,\alpha)}{\gamma_\mu(k)}x^k = e^{\alpha x^2}e_\mu(nx)(2\alpha x+n)$$

$$ii)\ \sum_{k=0}^{\infty}\frac{h^{\mu}_{k+2}(n,\alpha)}{\gamma_\mu(k)}x^k = e^{\alpha x^2}e_\mu(nx)(2\alpha^2x^2+4\alpha nx+2\alpha+n^2)+4\alpha\mu\ e^{\alpha x^2}e_\mu(-nx)$$

$$iii)\ \sum_{k=0}^{\infty}\frac{h^{\mu}_{k+3}(n,\alpha)}{\gamma_\mu(k)}x^k = e^{\alpha x^2}e_\mu(-nx)(8\alpha^2x\mu+4\alpha n\mu)+e^{\alpha x^2}e_\mu(nx)+e^{\alpha x^2}e_\mu(nx)(8\alpha^3x^3 + 12\alpha^2x^2n+12\alpha^2x+6\alpha xn^2+6an+n^3).$$

Proof. The proofs of above equalities are given in [3, 17]. □

Definition 2.2. *With the help of the Dunkl generalization of two variable Hermite polynomials given (9), we introduce the operators $D_n^H(f,x)$*
(14)

$$D_n^H(f,x)=\frac{1}{e^{\alpha x^2}e_\mu(nx)}\sum_{k=0}^{\infty}\frac{h^{\mu}_{k}(n,\alpha)}{\gamma_\mu(k)}x^k\frac{n^{k+2\mu\theta_k+\lambda+1}}{\Gamma(k+2\mu\theta_k+\lambda+1)}\int_0^\infty t^{k+2\mu\theta_k+\lambda}e^{-nt}f(\frac{nt+\gamma}{n+\beta})dt,$$

where $x\in[0,\infty)$,$\alpha\geq 0$,$\mu>-1/2$, $\lambda\geq 0$ and γ,β parameters satisfy the condition $0\leq\gamma\leq\beta$. We note that as a special case $\gamma=\beta=0$,then D_n^H reduces to S_n operator defined (13).

Lemma 2.3. *For the positive linear operator $D_n^H(f,x)$ given by (14), then the following results are satisfied.*

$$i)\ D_n^H(1,x)=1$$

$$ii)\ D_n^H(t,x)=\frac{nx}{n+\beta}+\frac{2\alpha x^2}{n+\beta}+\frac{\lambda+1}{n+\beta}+\frac{\gamma}{n+\beta}$$

$$iii)\ D_n^H(t^2,x)=\frac{4\alpha^2x^4}{(n+\beta)^2}+\frac{4n\alpha x^3}{(n+\beta)^2}+\frac{(10\alpha+n^2+4\lambda\alpha+4\gamma\alpha)x^2}{(n+\beta)^2}+\frac{(2\lambda n+4n+2\gamma n)}{(n+\beta)^2} +(\frac{2\mu n-4\alpha\mu x^2+4a\mu}{(n+\beta)^2})\frac{e_\mu(-nx)}{e_\mu(nx)}+\frac{\lambda^2+3\lambda+2}{(n+\beta)^2}+\frac{2\gamma\lambda+2\gamma}{(n+\beta)^2}.$$

Proof. **i)** By using the generating function (9), the relation (i) holds.

ii) With the help of definition of the Dunkl analogue of two-variable Hermite polynomials and the Gamma function, we get

$$D_n^H(f,x)=\frac{n}{(n+\beta)e^{\alpha x^2}e_\mu(nx)}\sum_{k=0}^{\infty}\frac{h^{\mu}_{k}(n,\alpha)}{\gamma_\mu(k)}x^k\frac{n^{k+2\mu\theta_k+\lambda+1}}{\Gamma(k+2\mu\theta_k+\lambda+1)}\frac{\Gamma(k+2\mu\theta_k+\lambda+2)}{n^{k+2\mu\theta_k+\lambda+1}} +\frac{\gamma}{(n+\beta)}\sum_{k=0}^{\infty}\frac{h^{\mu}_{k}(n,\alpha)}{\gamma_\mu(k)}x^k\frac{n^{k+2\mu\theta_k+\lambda+1}}{\Gamma(k+2\mu\theta_k+\lambda+1)}\frac{\Gamma(k+2\mu\theta_k+\lambda+1)}{n^{k+2\mu\theta_k+\lambda+1}}.$$

Also by using recurrence relation (7), we have

$$D_n^H(f,x)=\frac{n}{(n+\beta)}+\frac{2\alpha x^2}{(n+\beta)}+\frac{\lambda+1}{n+\beta}+\frac{\gamma}{n+\beta}$$

iii) Similarly for $f(t)=t^2$,

$$D_n^H(f,x)=\frac{1}{e^{\alpha x^2}e_\mu(nx)}\sum_{k=0}^{\infty}\frac{h^{\mu}_{k}(n,\alpha)}{\gamma_\mu(k)}x^k\frac{n^{k+2\mu\theta_k+\lambda+1}}{\Gamma(k+2\mu\theta_k+\lambda+1)} \times\left(n^2\int_0^\infty t^{k+2\mu\theta_k+\lambda+2}e^{-nt}dt+2n\gamma\int_0^\infty t^{k+2\mu\theta_{k+\lambda+1}}e^{-nt}dt+\gamma^2\int_0^\infty t^{k+2\mu\theta_{k+\lambda+2}}e^{-nt}dt\right),$$

where $\theta_{(k+1)}=\theta_k+(-1)^k$ and using the recurrence relation (7) completes the proof. □

Theorem 2.4. *The operator D_n^H and any uniformly continuous bounded function g on the interval $[0,\infty)$, we can give*

$$D_n^H\rightrightarrows g(x)$$

on each compact subset of $[0,\infty)$ when $n\to\infty$.

Proof. From the results obtained in Lemma 2

$$\lim_{n\to\infty} D_n^H(e^i;x) = x^i \qquad i = 0,1,2$$

holds. This convergence is uniformly in each compact subset of $[0,\infty)$. Then, applying the universal Korovkin type theorem gives the desired result. □

Lemma 2.5. *The operators $D_n^H(t;x)$ satisfiy the following results from Lemma 2.*

$$\Delta_1 = D_n^H((t-x);x) = \frac{nx}{(n+\beta)} + \frac{2\alpha x^2}{(n+\beta)^2} + \frac{(\lambda+1)}{(n+\beta)} + \frac{y}{(n+\beta)} - x,$$

$$\begin{aligned}\Delta_2 = D_n^H((t-x)^2;x) &= \frac{4\alpha^2}{(n+\beta)^2}x^4 + \frac{4\alpha n}{(n+\beta)}x^3 + \frac{(10\alpha+n^2+4\lambda\alpha+4\gamma\alpha)}{(n+\beta)^2}x^2 \\ &+ \frac{(2\lambda n+4n+2\gamma n)}{(n+\beta)^2}x + \frac{(2\mu nx-4\alpha\mu x^2+4\alpha\mu)}{(n+\beta^2}\frac{e_\mu(-nx)}{e_\mu(nx)} + \frac{\lambda^2+3\lambda+2}{(n+\beta)^2} + \frac{2\gamma\lambda+2\gamma}{(n+\beta)^2} \\ &- 2x\left(\frac{nx+2\alpha x^2+(\lambda+1)+\gamma}{(n+\beta)}\right) + x^2.\end{aligned}$$

Lemma 2.6. *The following inequality holds true for $h \in C_\infty^2[0,\infty]$*

$$D_n^H(h;x) - h(x) \le [\Delta_1 + \Delta_2] \|h\|_{C_\infty^2},$$

where Δ_1 and Δ_2 are central moments of operator D_n^H given by in Lemma 3.

Proof. □

Lagrange form of the remaining piece of the Taylor series

$$h(s) = h(x) + (s-x)h^{'}(x) + \frac{(s-x)^2}{2!}h^{''}(\sigma), \qquad \sigma \in (x,s).$$

Applying the operator D_n^H to both sides of this inequality and then using the linearity of the operator, we have

$$D_n^H(h;x) - h(x) = h^{'}(x)\Delta_1 + \frac{h^{''}(\sigma)}{2}\Delta_2$$

and, it yields

$$\left|D_n^H(h;x) - h(x)\right| \le \left\|h^{'}(x)\right\|\Delta_1 + \left\|h^{''}(x)\right\|\Delta_2 \le [\Delta_1+\Delta_2]\|h\|.$$

3. Special Convergence Results

In this part, we give some some definitions and rates of convergence results of the operators $D_n^H\ (f;x)$.

Let $g \in Lip_M(v)$ then the following inequality is hold

$$|g(s) - g(t)| \le M|s-t|^v,$$

where $s,t \in [0,\infty)$, $0 < \nu \le 1$, $M > 0$.

The modulus of continuity is defined by

$$w(g;\delta) := \sup_{\substack{s,t\in[0,\infty)\\ ||s-t||\le\delta}} |g(s) - g(t)|,$$

where $g \in \check{C}[0,\infty)$ which is the space of uniformly continuous on $[0,\infty)$.

Theorem 3.1. *If $f \in Lip_M(v)$, then we have*

$$\left|D_n^H(f;x) - f(x)\right| \le M(\Delta_2)^{v/2},$$

where $\Delta_2 = D_n^H((t-x)^2;x)$.

Proof. With the help of the definition of $f \in Lip_M(v)$ and linearity of operator D_n^H , we get

$$D_n^H(f(t)-f(x);x) \leq MD_n^H(|t-x|^v;x)$$
$$D_n^H(f(t);x)-f(x) \leq MD_n^H(|t-x|^v;x)$$
$$\left|D_n^H(f(t);x)-f(x)\right| \leq MD_n^H(|t-x|^v;x).$$

Let use the definition of operator

$$D_n^H(|t-x|;x) = D_n^H(f,x) = \frac{1}{e^{\alpha x^2}e_\mu(nx)}\sum_{k=0}^{\infty}\frac{h_k^\mu(n,\alpha)}{\gamma_\mu(k)}x^k\frac{n^{k+2\mu\theta_k+\lambda+1}}{\Gamma(k+2\mu\theta_k+\lambda+1)}$$
$$\times\left(\int_0^\infty t^{k+2\mu\theta_k+\lambda}e^{-nt}(\frac{nt+\gamma}{n+\beta}-x)^v dt\right)$$

and using Hölder's inequality for the above integral

$$\leq \frac{1}{e^{\alpha x^2}e_\mu(nx)}\sum_{k=0}^{\infty}\frac{h_k^\mu(n,\alpha)}{\gamma_\mu(k)}x^k\frac{n^{k+2\mu\theta_k+\lambda+1}}{\Gamma(k+2\mu\theta_k+\lambda+1)}$$
$$\times\left(\int_0^\infty t^{k+2\mu\theta_k+\lambda}e^{-nt}(\frac{nt+\gamma}{n+\beta}-x)^2 dt\right)^{v/2}\left(\int_0^\infty t^{k+2\mu\theta_k+\lambda}e^{-nt}\right)^{\frac{2-v}{2}}.$$

and then, we apply Hölder's inequality for sum

$$\leq\left(\sum_{k=0}^{\infty}\frac{1}{e^{\alpha x^2}e_\mu(nx)}\frac{h_k^\mu(n,\alpha)}{\gamma_\mu(k)}x^k\frac{n^{k+2\mu\theta_k+\lambda+1}}{\Gamma(k+2\mu\theta_k+\lambda+1)}\int_0^\infty t^{k+2\mu\theta_k+\lambda}e^{-nt}(\frac{nt+\gamma}{n+\beta}-x)^2 dt\right)^{v/2}$$
$$\times\left(\sum_{k=0}^{\infty}\frac{1}{e^{\alpha x^2}e_\mu(nx)}\frac{h_k^\mu(n,\alpha)}{\gamma_\mu(k)}x^k\int_0^\infty t^{k+2\mu\theta_k+\lambda}e^{-nt}\right)^{\frac{2-v}{2}}$$

by using the above inequalities, we get

$$D_n^H(|t-x|^v;x) \leq D_n^H(|t-x|^2;x)^{\frac{v}{2}}$$

from which, it follows

$$\left|D_n^H(f(t);x)-f(x)\right| \leq MD_n^H(|t-x|^v;x) \leq MD_n^H(|t-x|^2;x)^{\frac{v}{2}}$$

$$\left|D_n^H(f(t);x)-f(x)\right| \leq MMD_n^H(|t-x|^2;x)^{\frac{v}{2}}$$

□

Theorem 3.2. *$D_n^H((g(t);x)$ operators satisfiy the inequality*

$$\left|D_n^H(g(t);x)-g(x)\right| \leq 2w(g;\sqrt{\Delta_2}),$$

where $g \in C_B[0,\infty)$.

Proof. □

As a consequence of classical modulus of continuity and $g \in C_B[0,\infty)$

$$|g(t)-g(x)| \leq w(g;\delta)(\frac{|t-x|}{\delta}+1;x)$$

since D_n^H operator is also positive linear operator, we can write

$$\left|D_n^H(g;x)-g(x)\right| = \left|D_n^H(g(t)-g(x);x)\right| \leq D_n^H(|g(t)-g(x)|;x) \leq w(g;\delta)D_n^H(\frac{t-x}{\delta};x)$$
$$= w(g;\delta)(\frac{D_n^H(|t-x|;x)}{\delta}+1).$$

In addition to this inequaltiy, we use the definition of operator

$$D_n^H(|t-x|\,;x) = \frac{1}{e^{\alpha x^2}e_\mu(nx)}\sum_{k=0}^{\infty}\frac{h_k^\mu(n,\alpha)}{\gamma_\mu(k)}x^k\frac{n^{k+2\mu\theta_k+\lambda+1}}{\Gamma(k+2\mu\theta_k+\lambda+1)} \times\left(\int_0^\infty t^{k+2\mu\theta_k+\lambda}e^{-nt}(\frac{nt+\gamma}{n+\beta}-x)dt\right)$$

then by applying Cauchy Schwarz inequality

$$D_n^H(|t-x|\,;x) \leq \frac{1}{e^{\alpha x^2}e_\mu(nx)}\sum_{k=0}^{\infty}\frac{h_k^\mu(n,\alpha)}{\gamma_\mu(k)}x^k\frac{n^{k+2\mu\theta_k+\lambda+1}}{\Gamma(k+2\mu\theta_k+\lambda+1)} \times\left(\int_0^\infty t^{k+2\mu\theta_k+\lambda}e^{-nt}(\frac{nt+\gamma}{n+\beta}-x)^2dt\right)^{\frac{1}{2}}\left(\int_0^\infty t^{k+2\mu\theta_k+\lambda}e^{-nt}dt\right)^{\frac{1}{2}}.$$

and using Cauchy- Schwarz inequality for summation on right hand side

$$\leq\left(\sum_{k=0}^{\infty}\frac{1}{e^{\alpha x^2}e_\mu(nx)}\frac{h_k^\mu(n,\alpha)}{\gamma_\mu(k)}x^k\frac{n^{k+2\mu\theta_k+\lambda+1}}{\Gamma(k+2\mu\theta_k+\lambda+1)}\int_0^\infty t^{k+2\mu\theta_k+\lambda}e^{-nt}(\frac{nt+\gamma}{n+\beta}-x)^2dt\right)$$

$$\times\left(\sum_{k=0}^{\infty}\frac{1}{e^{\alpha x^2}e_\mu(nx)}\frac{h_k^\mu(n,\alpha)}{\gamma_\mu(k)}x^k\frac{n^{k+2\mu\theta_k+\lambda+1}}{\Gamma(k+2\mu\theta_k+\lambda+1)}\int_0^\infty t^{k+2\mu\theta_k+\lambda}e^{-nt}dt\right)^{\frac{1}{2}}$$

$$D_n^H(|t-x|\,;x)\leq\sqrt{\Delta_2}.$$

Thus, we can give the following inequaltiy

$$w(g;\delta)\left(\frac{D_n^H(|t-x|\,;x)}{\delta}+1\right)\leq w(g;\delta)\left(\frac{\sqrt{\Delta_2}}{\delta}+1\right)$$

so,

$$\left|D_n^H(g;x)-g(x)\right|\leq 2w(g;\delta),$$

where $\delta=\sqrt{\Delta_2}$. Hence, the proof complete.

Theorem 3.3. *For the operators $D_n^H(f;x)$, the following inequality holds*

$$\left|D_n^H(f;x)-g(x)\right|\leq 2K\left(g;\left(\sqrt{\Delta_1}+\frac{\sqrt{\Delta_2}}{2}\right)\right),$$

where Δ_1,Δ_2 are central moments of D_n^H.

Proof. Lagrange form of the remaining piece of the taylor series for $f\in C_B^2[0,\infty)$

$$f(t)-f(x)=f^{'}(x)(t-x)+f^{''}(c)\frac{(t-x)^2}{2}\qquad x\leq c\leq t$$

Then, by applying the operator $D_n^H(f;x)$ both side

$$\left|D_n^H(f;x)-g(x)\right|=\left|f^{'}(x)\right|D_n^H(|t-x|\,;x)+\frac{\left|f^{''}(c)\right|}{2}D_n^H(t-x)^2;x)$$

since $D_n^H((|t-x|\,;x)\leq\sqrt{\Delta_2}$, we get

$$\left|D_n^H(f;x)-g(x)\right|\leq\left|f^{'}(x)\right|\sqrt{\Delta_1}+\frac{\left|f^{''}(c)\right|}{2}\sqrt{\Delta_2}$$

$$\left|D_n^H(f;x)-g(x)\right|\leq\left\|f^{'}(x)\right\|_{C_B[0,\infty)}\sqrt{\Delta_1}+\frac{\left\|f^{''}(c)\right\|_{C_B[0,\infty)}}{2}\sqrt{\Delta_2}$$

$$\left|D_n^H(f;x)-g(x)\right|\leq\left(\sqrt{\Delta_1}+\frac{\sqrt{\Delta_2}}{2}\right)\|f\|_{C_B^2[0,\infty)}.$$

By using the above inequality, we can write

$$\begin{aligned}\left|D_n^H(f;x)-g(x)\right| &= \left|D_n^H(g;x)-D_n^H(f;x)+D_n^H(f;x)-f(x)+f(x)-g(x)\right|\\ &\leq D_n^H(|g-f|\,;x)+|g(x)-f(x)|+\left|D_n^H(f;x)-f(x)\right|.\end{aligned}$$

$$\left|D_n^H(f;x) - g(x)\right| \leq 2\left\|g-f\right\|_{c_B[0,\infty)} + 2\left\|f\right\|_{C_B^2[0,\infty)}\left(\sqrt{\Delta_1} + \frac{\sqrt{\Delta_2}}{2}\right). \tag{15}$$

The definition of Petre's K functional of the function$f \in C_B\ [0,\infty)$ is given as follows

$$K(f;\delta) := \inf_{g\in C_B^2[0,\infty)}\left\{\left\|f-g\right\|_{C_B[0,\infty)} + \delta\left\|g\right\|_{C_B[0,\infty)}\right\},$$

where for all $\delta > 0$. By taking infimum of both sides of (15) of for f, it yields

$$\left|D_n^H(f;x) - g(x)\right| \leq 2K\left(g;\sqrt{\Delta_1} + \frac{\sqrt{\Delta_2}}{2}\right)$$

Hence, the proof complete. □

References

[1] Altın A., Doğru O. and Tasdelen F., (2005), The generalization of Meyer-Konig and Zeller operators by generating functions, J. Math. Anal. Appl. 312 (1),181-194, 2005.

[2] Atakut C., Ispir N.,(2002), Approximation by modified Szász–Mirakjan operators on weighted spaces, Proc. Indian Acad. Sci. Math. 112 (2002), 571–578.

[3] Aktas R., Cekim B. and Tasdelen F.,A Dunkl Analogue of Operators Including TwoVariable Hermite polynomials, Bull. Malays. Math. Sci. Soc., 42 (2019), 2795–2805.

[4] Becker M., Global Approximation Theorems for Szász-Mirakjan and Baskakov Operators in Polynomial Weight Spaces Indiana University Mathematics Journal Vol. 27 No.1,pp. 127-142.

[5] Ben Cheikh Y., Gaied M., Dunkl–Appell d-orthogonal polynomials. Integral Transforms and Special Functions, 18(8) (2007), 581-597.

[6] Cekim B., Aktas R., Tasdelen F., A Dunkl-Gamma Type Operator in Terms of Generalization of Two-Variable Hermite Polynomials, Indian J Pure Appl Math 53, 727–735 (2022).

[7] Krech G. A note on some positive linear operators associated with the Hermite polynomials, Carpathian J. Math., 32 (1) (2016), 71–77.

[8] Korovkin P. On convergence of linear positive operators in the space of continuous functions (Russian), Doklady Akad. Nauk. SSSR (NS) 90 (1953), 961–964.

[9] Mazhar S., Totik V., Approximation by modified Szász operators, Acta Sci. Math., 49(1-4) (1985), 257–269.

[10] Mirakjan G. M., Approximation of continuous functions with the aid of polynomials, In Dokl. Acad. Nauk. SSSR., 31 (1941), 201–205

[11] Dogru O., Özarslan M.A. and Tasdelen F., On positive operators involving a certain class of generating functions, Stud. Sci. Math. Hungar. 41 (4), 415-429, 2004.

[12] Aktaş R., Çekim B. and Taşdelen F., A Kantorovich-Stancu type generalization of Szász operators including Brenke-type polynomials, J. Func. Spaces 2013, 9 pages, 2013.

[13] Rosenblum M., Generalized Hermite polynomials and the Bose-like oscillator calculus, Oper. Theory: Adv. Appl. 73 (1994), 369-396.

[14] Szász O., Generalization of the Bernstein polynomials to the infinite interval, J. Res. Nat. Bur. Stand., 45(3) (1950), 239–245. https://doi.org/10.6028/jres.045.024

[15] Bernstein S.N., Demonstration du theoreme de Weierstrass fondee sur le calcul des probabilites, Commun. Soc. Math. Kharkow, 2(13) (2012) 1–2.

[16] Sucu S., Dunkl analogue of Szász operators, Appl. Math. Comput. 244 (2014), 42-48.

[17] Tasdelen F., Soylemez D., Aktas R., Dunkl-Gamma type operators including Appell polynomials, Complex Analysis and Operator Theory, 13 (2019), 3359–3371.

[18] Wafi A., Rao N., Szász –gamma operators based on Dunkl analogue, Iran J Sci Technol Trans Sci., 43(2019),213–223.

Zeynep Er[1*]

Department of Mathematics, Faculty of Science, Ankara University, Tandogan, 06100 Ankara, Türkiye

Harun Karsli[2]

Department of Mathematics, Faculty of Science and Arts, Abant Izzet Baysal University, Tandogan, 14280 Bolu, Türkiye

Rabia Aktaş Karaman[3]

Department of Mathematics, Faculty of Science, Ankara University, Tandogan, 06100 Ankara, Türkiye

Fatma Taşdelen Yeşildal[4]

Department of Mathematics, Faculty of Science, Ankara University, Tandogan, 06100 Ankara, Türkiye

Email(s): zsakartepe@ankara.edu.tr[1*], karsli_h@ibu.edu.tr[2], raktas@science.ankara.edu.tr[3] tasdelen@science.ankara.edu.tr[4]

A mathematical model for the effects of wavelets and the analysis of neural network operators described using wavelets

Harun Karslı

ABSTRACT. It is very well-known that wavelets have great advantage of being able to separate and identify fine details in a signal or a function. One of the main advantages of wavelets compared to the Fourier analysis and its related theories is that they offer simultaneous localization in the time and frequency domain. The second main advantage of wavelets is that they are computationally very fast and detailed when using wavelet expansions and transformations. In the present study we will deal with the linear approximation operators constructed by compactly supported Daubechies wavelets. In details, we will reconstruct neural network operators, where location and time are very important and effective, with the help of wavelets, and we will examine and analyse various properties of the wavelet type extension of the neural network operators.

2020 MATHEMATICS SUBJECT CLASSIFICATIONS: 41A35, 65T60, 41A05, 47A58, 41A25

KEYWORDS: wavelets, neural network operators, approximation.

1. INTRODUCTION

It is known that Artificial Neural Networks (ANN) are defined as a mathematical model that allows the brain's behavior and abilities to be reproduced. In general, Neural Network is a family of parametrized functions, namely by a mathematical point of view it is a multiple composition of some special functions called activations. It is important to point out that, this approach motivated by the fundamental density theorem due to George Cybenko given in [23].

In 1989, Cybenko [23] gave an answer to the superposition problem on $C[a,b]$ with his famous density theorem, which states that every continuous function defined on $[a,b]$ can be approximated by a sequence constructed by a linear combination of sigmoidal functions. In other words, Cybenko confirmed that a neural network with solely one hidden-layer is capable of always approximating to a continuous function. The main advantage of the above theory lies on its connection with the Artificial Neural Networks, Learning Theory and their applications to Approximation Theory, see e.g. [4], [22] and [33].

Based on the idea developed by Cybenko, the theory of the mathematical models of the neural network (NN) operators arise since 1992 with the pioneer work of Cardaliaguet and Euvrard [12], and then in the next years, they have been largely studied by several authors under different aspects. Especially in the last two decades, there are many new version of artificial neural networks has been introduced and widely studied.

In particular, in 1997 Anastassiou [2] pointed out and obtained that the compactly supported bell-shaped functions used in the Density Theorem of Cybenko and the Cardaliaguet and Euvrard (NN) operators can be obtained from sigmoidal functions used effectively in Artificial Neural Networks, serious relations have emerged between the Cybenko convergence theorem and the Theory of Approximations. Thanks to this fundamental work of Anastassiou in 1997, the techniques and theorems of the Theory of Approximations began to be used in Artificial Neural Networks.

Owing to the work of the famous mathematician G. Anastassiou in 1997, see [2], the theory of neural network(NN) operators has been introduced by Anastassiou as an extension of the bell-shaped operators studied by Cardaliaguet and Euvrard in [12].

Especially in the last two decades, based on the idea developed by Anastassiou [2] and afterwards the continuous works on these operators and some of their modifications of D.

Costarelli, R. Spigler, C. Bardaro, G. Vinti and their research group (RITA network) from Perugia, such as Kantorovich and Durrmeyer forms have been of great importance in the development of mathematical models for signal and image recovering. And hence the approximation problem was proved by using NN operators and some of its different forms in some function spaces, see e.g., [3]-[6], [10]-[11], [14]-[21], [7], [9] and [13]). Since the classical neural network operators cannot be used for $L^p(\mathbb{R})(1 \leq p < \infty)$ approximation, to obtain some positive results for these functions, their Kantorovich and Durrmeyer type modifications are considered.

Especially, in the paper [15], Costarelli and Spigler introduced and investigated a kind of neural network operators for absolutely continuous functions, by considering the relations between sigmoidal functions, multiresolution analysis and the scaling functions. Moreover they also state some intersting results related with the rate of convergence in the set $AC[a,b]$.

The goal of this study is to find a positive solution to the approximation (or superposition) problem for operators in some general function spaces, by using the effects and relations between different function spaces provided by the wavelets . In other words, we will propose a generalization and extension of the theory of interpolation to operators by introducing an integral operator, called wavelet type operators (see [1], [26], [31], [32] and [35]). The new operators are more flexible than the previous ones and they are at least a natural extensions of the classical NN operators, Kantorovich and Durrmeyer type modifications.

The basis used in this construction of the new neural network operators are the Cybenko density theorem and wavelets.

Similar constructions and investigations can be found in the very recent papers of the author [27]-[30].

It is very well-known that wavelets and wavelet expansions have the great advantage of being able to separate and identify fine details in a signal or a function.

One of the main advantages of wavelets compared to the Fourier analysis and its related theories is that they offer simultaneous localization in the time and frequency domain. The second main advantage of wavelets is that they are computationally very fast and detailed when using wavelet expansions and transformations.

Unlike the Fourier analysis, wavelets tell us about the frequencies present as well as the time in which these frequencies were observed.

So, wavelets are a better way of analyzing especially the dynamic signals because they have a relatively higher resolution in both time and frequency domain.

Moreover, from the definitions and properties of the wavelet bases, one can use wavelet type operators for approximation problem in L^p spaces.

Since wavelets have many advantages for approximating in L^p spaces, potential applications in machine learning and neural networks, the future directions of this work are to try to adapt what has been accomplished in wavelets to these spaces and theories (see [5] and [22]).

In Sect. 2, we recall the definition of the NN operators and wavelets, and all their main properties which are useful in order to prove the quantitative estimates, together with some examples of activation functions, while in Sect. 3 the main results of the paper have been established. In the final part of this study, we will provide also some graphical examples and comparisons between the convergence of the NN operators obtained different kind of wavelets.

2. Preliminaries and auxiliary results

In this section we shall recall some notation and background material of the theory of Neural Networks and the theory of wavelet, especially Daubechies' compactly supported wavelets ([24], [25]), used throughout this paper.

We denote by $C[a,b], B[a,b]$ and $L_\infty(\mathbb{R})$ the sets of continuous, bounded and essentially bounded functions with their usual norms, respectively.

Definition 2.1. ***(Centered bell-shaped function)*** *A function $b : \mathbb{R} \to \mathbb{R}$ is said to be centered bell-shaped if b belongs to L^1 and its integral is nonzero, if it is nondecreasing on $(-\infty, 0)$ and nonincreasing on $[0, +\infty)$.*

Definition 2.2. ***(Sigmoidal Function)*** *Let $\sigma : \mathbb{R} \to \mathbb{R}$ be a measurable function satisfies*

$$\lim_{x \to -\infty} \sigma(x) = 0 \qquad \text{and} \qquad \lim_{x \to \infty} \sigma(x) = 1,$$

then it is called sigmoidal function.

As an example of a sigmoidal and its corresponding bell-shaped function, we can consider the ramp function, which is very useful and important for the neural network.

Definition 2.3. ***(Ramp Function)*** *A ramp function is a special sigmoidal function defined as*

$$R(x) = \begin{cases} 0 & , \quad x \leq -1/2 \\ x + 1/2 & , \quad -1/2 < x < 1/2 \\ 1 & , \quad x \geq 1/2 \end{cases}.$$

Clearly, a bell-shaped function can be define by using Sigmoidal (or Ramp) function, namely

$$b_\sigma(x) = \frac{\sigma(x+1) - \sigma(x-1)}{2},$$

and

$$b_R(x) = R(x + 1/2) - R(x - 1/2).$$

Moreover, $b_R : \mathbb{R} \to \mathbb{R}$ is a bell-shaped kernel function obtained by ramp function $R(x)$ that satisfies following assumptions:

b_R is a continuous function on $\mathbb{R}$,

$$b_R \in L^1(\mathbb{R}), \ \sum_{k=0}^{n} b_R(u-k) = 1 \text{ for every } u \in \mathbb{R}, \tag{1}$$

and

$$O_R := \sup_{u \in \mathbb{R}} \sum_{k=0}^{n} b_R(u-k) < \infty \tag{2}$$

where the convergence of the series (2) is uniform on each compact subintervals of $\mathbb{R}$.

Some other sigmoidal functions are Gompertz function, Logistic function, Error function and Hyperbolic tangent functions, etc.

It is important to note that, the theory of sigmoidal functions are not new. At 1838, the logistic function was introduced by Pierre François Verhulst [36]-[37], who applied it to human population dynamics. Verhulst derived his logistic equation to describe the mechanism of the self-limiting growth of a biological population.

Since then the logistic functions have many applications in many research areas, including biology, ecology, population dynamics, chemistry, demography, economics, geoscience, mathematical psychology, probability, sociology, political science, nancial mathematics, statistics.

Now we will give some definitions of the Neural Network (NN) Operators.

Definition 2.4. ***(Cardaliaguet and Euvrard (NN) Operators)***

Let $f : \mathbb{R} \to \mathbb{R}$ be a continuous and bounded function and $n \in \mathbb{N}^+$, the Cardaliaguet and Euvrard (NN) Operators are defined as.

$$(F_n f)(x) = \sum_{k=-n^2}^{n^2} \frac{f\left(\frac{k}{n}\right)}{B n^\alpha} b(n^{-\alpha}(nx - k),$$

where $0 < \alpha < 1, b$ is a bell-shaped function with compact support $\subset [-T, T]$, and

$$B := \int_{-T}^{T} b(x) dx$$

The Cardaliaguet and Euvrard (NN) Operators and Its different modifications were intensively studied by Anastassiou, Spiegler, Costarelli, Vinti.

In [2], Anastassiou defined the Cardaliaguet and Euvrard (NN) Operators as follows;

Definition 2.5. *Let $f:\mathbb{R}\to\mathbb{R}$ be a continuous and bounded function, $T>0, n\in\mathbb{N}^{+}$ and $n\geq\max\{T+|x|, T^{-1/\alpha}\}$, the Cardaliaguet and Euvrard (NN) Operators are given by*

$$(F_n f)(x)=\sum_{k=\lceil nx-Tn^{\alpha}\rceil}^{\lfloor nx+Tn^{\alpha}\rfloor}\frac{f\left(\frac{k}{n}\right)}{Bn^{\alpha}}b(n^{-\alpha}(nx-k),$$

where again $0<\alpha<1, b$ is a bell-shaped function with compact support $\subset[-T,T]$, and

$$B:=\int_{-T}^{T}b(x)dx.$$

As a special case of the operators defined on $[a,b]$ and $[0,1]$, we have the following type Neural Network (NN) Operators, respectively.

Definition 2.6. ***(Neural Network (NN) Operators).*** *Let $f:[a,b]\to\mathbb{R}$ be a bounded function, and $n\in\mathbb{N}^{+}$ such that $\lceil na\rceil\leq\lfloor nb\rfloor$. The positive linear neural network operators activated by the ramp function $R(x)$, are defined as.*

$$(F_n f)(x)=\frac{\sum_{k=0}^{n}f\left(a+k\frac{b-a}{n}\right)b_R\left(\frac{n(x-a-k\frac{b-a}{n})}{b-a}\right)}{\sum_{k=0}^{n}b_R\left(\frac{n(x-a-k\frac{b-a}{n})}{b-a}\right)},\quad x\in[a,b],$$

and

Definition 2.7. ***(Neural Network (NN) Operators).*** *Let $f:[0,1]\to\mathbb{R}$ be a bounded function, and $n\in\mathbb{N}^{+}$. The positive linear neural network operators activated by the ramp function R, are defined as.*

$$(N_n f)(x)=\frac{\sum_{k=0}^{n}f\left(\frac{k}{n}\right)b_R(nx-k)}{\sum_{k=0}^{n}b_R(nx-k)},\quad x\in[0,1],\tag{3}$$

where b_R is the bell-shaped function obtained by the ramp function R.

Now, we will give some informations about wavelets, multiresolution analysis and its theories, which will be useful for the remain part of this work.

Definition 2.8. ***(Scale Function)*** *A scale function defined on the interval $[0,1)$ as*

$$\phi(t)=\begin{cases}1 & , \quad 0\leq x<1\\ 0 & , \quad e.w.\end{cases}.$$

Clearly, a scale function can be also define by using Heaviside unit step function, namely

$$H(x)=\begin{cases}1 & , \quad x\geq 0\\ 0 & , \quad x<0\end{cases},$$

and hence

$$\phi(x)=H(x)-H(x-1).$$

A multiresolution analysis (MRA) is a sequence $(Vj)_{j\in Z}$ of closed subspaces of $L^2(\mathbb{R})$ such that the following hold:

i) V_j is a set of all $f\in L^2(\mathbb{R})$ which are constant on 2^{-j} length intervals and

$$...\subset V_{-2}\subset V_{-1}\subset V_0\subset V_1\subset...\subset V_j\subset V_{j+1}\subset..\subset L_2(\mathbb{R}),$$

$$\overline{\bigcup_j V_j}=L_2(\mathbb{R}),$$

ii)

$$\begin{aligned}\forall j,k &\in \mathbb{Z},\ f(x)\in V_j \Leftrightarrow f(2x)\in V_{j+1},\\ \forall k &\in \mathbb{Z},\ f(x)\in V_0 \Leftrightarrow f(x-k)\in V_0\\ \forall j,k &\in \mathbb{Z},\ f(x)\in V_j \Leftrightarrow f(x-2^{-j}k)\in V_j,\end{aligned}$$

and

iii)

$$\bigcap_j V_j = \{0\}\,.$$

Since $V_j \subset V_{j+1}$ ($\forall j \in \mathbb{Z}$) then there exist subspaces W_j of $L_2(\mathbb{R})$ satisfying

$$\begin{gathered}V_1 = V_0 \oplus W_0,\\ V_2 = V_1 \oplus W_1,\\ \ldots\\ V_{j+1} = V_j \oplus W_j,\end{gathered}$$

and

$$L_2(\mathbb{R}) = \ldots W_{-2}\oplus W_{-1}\oplus W_0 \oplus \ldots \oplus W_j \oplus W_{j+1}\oplus \ldots$$

Definition 2.9. *(**Wavelet**) A wavelet is a small wave which oscillates and decays in the time domain. A wavelet basis set starts with two orthogonal functions: the scaling function or father wavelet $\phi(t)$ and the wavelet function or mother wavelet $\vartheta(t)$. By scaling and translation of these two orthogonal functions we obtain a complete basis set. The scaling and wavelet functions, respectively, satisfy*

$$\int_{-\infty}^{\infty}\phi(t)dt = 1,\ \int_{-\infty}^{\infty}\vartheta(t)dt = 0.$$

These two functions have finite energy, namely $\phi, \vartheta \in L^2(\mathbb{R})$, and orthogonal.

In general, the wavelets refers to the set of family of orthonormal functions of the form

$$\begin{aligned}\phi_{a,b}(t) &= \frac{1}{\sqrt{a}}\phi\left(\frac{t-b}{a}\right), \quad a>0, b\in\mathbb{R},\\ \vartheta_{a,b}(t) &= \frac{1}{\sqrt{a}}\vartheta\left(\frac{t-b}{a}\right), \quad a>0, b\in\mathbb{R},\end{aligned}$$

where ϕ and ϑ are the basic, father and mother wavelets, respectively.

Haar Wavelet: The simplest wavelet is known as the Haar Wavelet defined as;

$$\vartheta(x) = \begin{cases} 1 & , \quad 0 \le x < \frac{1}{2}\\ -1 & , \quad \frac{1}{2} \le x < 1\\ 0 & , \quad e.w.\end{cases}$$

with the scaling function

$$\phi(t) = \begin{cases} 1 & , \quad 0 \le x < 1\\ 0 & , \quad e.w.\end{cases}.$$

Clearly, Haar wavelets constitutes an orthonormal system for the space of square-integrable functions on the real line. Since Haar wavelet is not continuous and therefore not differentiable, it is suitable for representing discrete signals not for representing smooth signals or functions.

In the present study, we consider orthonormal bases of wavelets in $L^2(\mathbb{R})$, and assume that there is a scaling function (father wavelet) $\phi(t)$ whose whose translates $\{\phi(t-n)\}$ are orthogonal and the mother wavelet $\vartheta(t)$ based on the father wavelet $\phi(t)$ gives rise to the orthonormal basis $\vartheta_{j,k}(t)$ of $L^2(\mathbb{R})$, where

$$\vartheta_{j,k}(t) = 2^{j/2}\vartheta(2^j t - k). \tag{4}$$

Hence, by using a multiresolution analysis (MRA), each $f \in L^2(\mathbb{R})$ has the following representation

$$f(x) = \sum_{j\in\mathbb{Z}} \sum_{k\in\mathbb{Z}} b_{j,k} \vartheta_{j,k}(x),$$

called wavelet expansion of $f \in L^2(\mathbb{R})$, where $b_{j,k}$ are wavelet coefficients defined by

$$b_{j,k} = \langle f(x), \vartheta_{j,k}(x) \rangle = 2^{j/2} \int_{\mathbb{R}} f(x) \overline{\vartheta(2^j x - k)} dx.$$

Some convergence results about wavelet expansions, please see [31], [35] and [32].

Some of the special cases of a and b, one can obtain different type of wavelets.

Franklin system: Let $k, j \in \mathbb{Z}$. If one choose $a = 2^{-j}$ and $b = k2^{-j}$, then one can obtain the Franklin system, which is an orthonormal basis of $L^2[0,1]$.

Strömberg wavelet: Even though the Haar wavelet was earlier known to be an orthonormal wavelet, Strömberg wavelet was the first smooth orthonormal wavelet to be discovered. Namely, J.O. Strömberg constructed the orthonormal basis of the form (4) by using a function $\vartheta \in C^m$ for an abitrary nonnegative integer m, which is a complete orthonormal system in the space of square integrable functions over $\mathbb{R}$.

The wavelet analysis procedure is to adopt a wavelet prototype function, called an analyzing wavelet (father wavelet) or mother wavelet.

Definition 2.10. *([24], [25]) **(Compactly supported Daubechies Wavelet)***

Owing to the above definitions, first of all, we introduce the NN operators by using the compactly supported Daubechies wavelets considered in this paper.

Let us assume that the scale function (or father wavelets) $\psi \in L_\infty(\mathbb{R})$ and satisfies:

a) ψ is a compactly supported, namely there is a real constant $\lambda > 0$ such that supp $\psi \subset [0, \lambda]$,

b) $\int_{-\infty}^{\infty} \psi(x)dx = 1,$

c) the first N moments of the father wavelet ψ satisfy

$$\int_{-\infty}^{\infty} x^j \psi(x) dx = 0, \qquad j = 1, ..., N.$$

Note 1. Actually, Daubechies wavelets have strong relations with the properties of continuity and differentiability. Namely, for an arbitrary fixed integer $N \geq 1$, compactly supported Daubechies wavelet ψ is supported with $[0, 2N-1]$, in addition there exists a constant $r > 0$ such that for $N \geq 2$, $\psi \in C^{rN}(\mathbb{R})$ and to have a given number of vanishing moments.

In particular, when $N = 1$, then the first Daubechies wavelet ψ will be the classical Haar basis. As N increases, the regularity of the wavelets increase (see [24], [25]).

This means that if we want to use Daubechies wavelets to reconstruct a function, it is more convenient to choose or construct wavelets based on the continuity or differentiability properties of the given function (please see Haar wavelet).

Now, we will consider neural network operators activated by the ramp function R, where location and time are very important and effective, with the help of wavelets.

Moreover we will examine and analyse various properties of the wavelet type extension of the neural network operators.

Definition 2.11. ***(Wavelet type Neural Network (NN) Operators).*** *Let $f:[0,1]\to\mathbb{R}$ be a bounded measurable function, $n\in\mathbb{N}^{+}$. and let $\psi\in L_{\infty}(\mathbb{R})$ be a father wavelet satisfies* ***a)-c)****. Then the wavelet type Neural Network (NN) operators activated by the ramp function R, constructed by using the compactly supported Daubechies wavelets, are defined by:*

$$(WN_n f)(t) = n\frac{\sum\limits_{k=0}^{n} b_R(nt-k)\int\limits_{\mathbb{R}} f(x)\,\psi(nx-k)dx}{\sum\limits_{k=0}^{n} b_R(nt-k)}, \quad t\in[0,1], \tag{5}$$

where b_R is the bell-shaped function obtained by the ramp function R.

Remark 1. *If we choose the father wavelet ψ as the Haar scaling function, namely $\psi(x)=\chi_{[0,1]}(x)$, then clearly our wavelet type operators reduce to the Kantorovich form of the Neural Network (NN) operators. Indeed;*

$$\begin{aligned}(WN_n f)(t) &= n\frac{\sum\limits_{k=0}^{n} b_R(nt-k)\int\limits_{\mathbb{R}} f(x)\,\psi(nx-k)dx}{\sum\limits_{k=0}^{n} b_R(nt-k)}\\ &= \frac{\sum\limits_{k=0}^{n} b_R(nt-k)\int\limits_{0}^{1} f\left(\frac{u+k}{n}\right)\psi(u)dx}{\sum\limits_{k=0}^{n} b_R(nt-k)}.\end{aligned}$$

This means that our operators constructed by wavelets are a natural extension of the Kantorovich type of the NN operators and also its Durrmeyer type operators.

Remark 2. *Moreover, the central moments of the wavelet type NN operators (5) are the same as of the classical NN operators (3). Indeed, we get*

$$\begin{aligned}(WN_n(x-t)^{\beta})(t) &= n\frac{\sum\limits_{k=0}^{n} b_R(nt-k)\int\limits_{\mathbb{R}}(x-t)^{\beta}\,\psi(nx-k)dx}{\sum\limits_{k=0}^{n} b_R(nt-k)}\\ &= \frac{\sum\limits_{k=0}^{n} b_R(nt-k)\int\limits_{\mathbb{R}}\left(\frac{u+k}{n}-t\right)^{\beta}\psi(u)du}{\sum\limits_{k=0}^{n} b_R(nt-k)}\\ &= \frac{1}{n^{\beta}}\frac{\sum\limits_{k=0}^{n} b_R(nt-k)\int\limits_{\mathbb{R}}(u+k-nt)^{\beta}\,\psi(u)du}{\sum\limits_{k=0}^{n} b_R(nt-k)}\\ &= \frac{1}{n^{\beta}}\frac{\sum\limits_{k=0}^{n} b_R(nt-k)\int\limits_{\mathbb{R}}\left[\sum\limits_{i=0}^{\beta}\binom{\beta}{i}u^{i}(nt-k)^{\beta-i}\right]\psi(u)du}{\sum\limits_{k=0}^{n} b_R(nt-k)}\end{aligned}$$

Again by the properties of the compactly supported Daubechies wavelets, namely **c)** and **b)**, we get

$$\begin{aligned}(WN_n(x-t)^{\beta})(t) &= \frac{1}{n^{\beta}}\frac{\sum\limits_{k=0}^{n} b_R(nt-k)(k-nt)^{\beta}}{\sum\limits_{k=0}^{n} b_R(nt-k)}\\ &= (N_n(x-t)^{\beta})(t).\end{aligned}$$

Throughout this work, for every $u \in \mathbb{R}$ and for some $\beta > 0$, we assume that the algebraic and discrete absolute moment of order β are given by, i.e.,

$$m_\beta(b_R) := \sup_{u\in\mathbb{R}} \sum_{k=0}^{n} b_R\,(u-k)\,(u-k)^\beta\,,$$

and

$$M_\beta(b_R) := \sup_{u\in\mathbb{R}} \sum_{k=0}^{n} b_R\,(u-k)\,|u-k|^\beta < \infty.$$

3. Fundamental Properties and Main Results

We now introduce some notations and structural hypotheses, which will be fundamental in proving our convergence theorems. This section provides the main approximation results of the paper. We have the followings.

We are now ready to establish one of the first main results of this study, which gives a strong relation between NN operators (3) and our new operators (5) constructed by wavelets.:

Theorem 3.1. *Let $f \in B[0,1]$ and let $\psi \in L_\infty(R)$ be a father wavelet satisfies* ***a)-c)****. Then the moments of wavelet type NN operators, constructed by using the compactly supported Daubechies wavelets (5) and the NN operators (3) are the same, namely*

$$(WN_n x^s)(t) = (N_n x^s)(t), \quad s = 0,1,...,K$$

holds true.

Proof. In view of the definition of the operator (5), we have

$$\begin{aligned}
(WN_n x^s)(t) &= n\frac{\sum\limits_{k=0}^{n}\left[\int\limits_{\mathbb{R}} x^s\psi(nx-k)dx\right] b_R(nt-k)}{\sum\limits_{k=0}^{n} b_R(nt-k)} \\
&= \frac{\sum\limits_{k=0}^{n} b_R(nt-k)\int\limits_{\mathbb{R}}\left(\frac{u+k}{n}\right)^s\psi(u)du}{\sum\limits_{k=0}^{n} b_R(nt-k)} \\
&= \frac{1}{n^s}\frac{\sum\limits_{k=0}^{n} b_R(nt-k)\int\limits_{\mathbb{R}}(u+k)^s\psi(u)du}{\sum\limits_{k=0}^{n} b_R(nt-k)} \\
&= \frac{1}{n^s}\frac{\sum\limits_{k=0}^{n} b_R(nt-k)\int\limits_{\mathbb{R}}\left[\sum\limits_{i=0}^{s}\begin{pmatrix} s \\ i\end{pmatrix} u^i k^{s-i}\right]\psi(u)du}{\sum\limits_{k=0}^{n} b_R(nt-k)}.
\end{aligned}$$

In view of **c)**, one has for $i \neq 0$

$$\int\limits_{\mathbb{R}}\left[\sum_{i=0}^{s}\begin{pmatrix} s \\ i\end{pmatrix} u^i k^{s-i}\right]\psi(u)du = 0$$

and for $i = 0$ and from **b)** we get

$$\begin{aligned}
(WN_n x^s)(t) &= \frac{1}{n^s}\frac{\sum\limits_{k=0}^{n} b_R(nt-k)\int\limits_{\mathbb{R}} k^s\psi(u)du}{\sum\limits_{k=0}^{n} b_R(nt-k)} \\
&= \frac{\sum\limits_{k=0}^{n}\frac{k^s}{n^s} b_R(nt-k)}{\sum\limits_{k=0}^{n} b_R(nt-k)} \\
&= (N_n x^s)(t).
\end{aligned}$$

□

Theorem 3.2. *Let $f \in B[0,1]$ be a measurable function and let $\psi \in L_\infty(\mathbb{R})$ be a father wavelet satisfies* ***a)-c).*** *Then*

$$\lim_{n\to\infty} (WN_n f)(t_0) = f(t_0)$$

holds true at each point t_0 of continuity of f.

Proof. Since t_0 is a continuity point of f, then clearly

$$|f(t) - f(t_0)| < \epsilon$$

holds true when $|t - t_0| < \delta$, and

$$|f(t) - f(t_0)| \leq 2\,\|f\|_{B[0,1]}$$

holds true, when $|t - t_0| \geq \delta$.

So, in view of the definition of the operator (5), one has

$$\begin{aligned}
(WS_n f)(t_0) - f(t_0) &= n\frac{\sum\limits_{k=0}^{n} b_R(nt_0-k)\int\limits_{\mathbb{R}} f(x)\,\psi(nx-k)dx}{\sum\limits_{k=0}^{n} b_R(nt_0-k)} - f(t_0) \\
&= \frac{\sum\limits_{k=0}^{n} b_R(nt_0-k)\int\limits_{\mathbb{R}} f\left(\frac{u+k}{n}\right)\psi(u)du}{\sum\limits_{k=0}^{n} b_R(nt_0-k)} - f(t_0).
\end{aligned}$$

By Theorem 1, we know that

$$(WN_n 1)(t) = (N_n 1)(t) = 1. \tag{6}$$

Hence we can write

$$\begin{aligned}
|(WS_n f)(t_0) - f(t_0)| &= \left|\frac{\sum\limits_{k=0}^{n} b_R(nt_0-k)\int\limits_{\mathbb{R}} \left[f\left(\frac{u+k}{n}\right) - f(t_0)\right]\psi(u)du}{\sum\limits_{k=0}^{n} b_R(nt_0-k)}\right| \\
&\leq \frac{\sum\limits_{k=0}^{n} b_R(nt_0-k)\int\limits_{\mathbb{R}} \left|f\left(\frac{u+k}{n}\right) - f(t_0)\right| |\psi(u)|\,du}{\sum\limits_{k=0}^{n} b_R(nt_0-k)}
\end{aligned}$$

Let us divide the last term into two parts as;

$$|(WN_n f)(t_0) - f(t_0)| \leq P_1 + P_2,$$

where

$$P_1 \leq \frac{\sum\limits_{k=0}^{n} b_R(nt_0-k)\int\limits_{\left|\frac{u+k}{n}-t_0\right|<\delta} \left|f\left(\frac{u+k}{n}\right) - f(t_0)\right| |\psi(u)|\,du}{\sum\limits_{k=0}^{n} b_R(nt_0-k)}$$

and

$$P_2 \leq \frac{\sum\limits_{k=0}^{n} b_R(nt_0-k)\int\limits_{\left|\frac{u+k}{n}-t_0\right|\geq\delta} \left|f\left(\frac{u+k}{n}\right) - f(t_0)\right| |\psi(u)|\,du}{\sum\limits_{k=0}^{n} b_R(nt_0-k)}$$

Hence one has

$$\begin{aligned}
P_1 &= \frac{\sum\limits_{k=0}^{n} b_R(nt_0-k)\int\limits_{\left|\frac{u+k}{n}-t_0\right|<\delta} \left|f\left(\frac{u+k}{n}\right) - f(t_0)\right| |\psi(u)|\,du}{\sum\limits_{k=0}^{n} b_R(nt_0-k)} \\
&\leq O_R\epsilon\,\|\psi\|_\infty,
\end{aligned}$$

and

$$
\begin{aligned}
P_2 &= \frac{\sum\limits_{k=0}^{n} b_R\left(nt_0-k\right) \int\limits_{\left|\frac{u+k}{n}-t_0\right|\geq\delta} \left|f\left(\frac{u+k}{n}\right)-f(t_0)\right| \left|\psi(u)\right| du}{\sum\limits_{k=0}^{n} b_R\left(nt_0-k\right)} \\
&\leq 2\left\|f\right\|_{B[0,1]} \frac{\sum\limits_{k=0}^{n} b_R\left(nt_0-k\right) \int\limits_{\left|\frac{u+k}{n}-t_0\right|\geq\delta} \left|\psi(u)\right| du}{\sum\limits_{k=0}^{n} b_R\left(nt_0-k\right)} \\
&\leq 2\left\|f\right\|_{B[0,1]} \frac{M_2(b_R)}{\delta^2 n^2} \left\|\psi\right\|_{\infty} = O(n^{-2}).
\end{aligned}
$$

Collecting these estimates we have

$$\lim_{n\to\infty} \left(WN_n f\right)(t_0) = f(t_0).$$

This completes the proof. □

Theorem 3.3. *Let $f \in C[0,1]$ and let $\psi \in L_\infty(\mathbb{R})$ be a father wavelet satisfies **a)-c)**. Then*

$$\left|\left(WN_n f\right)(x) - f(x)\right| \leq 4\omega\left(f; 1/n\right)$$

holds true.

Corollary 3.4. *The same arguments of Theorem 2 apply to the case when $f \in C[0,1]$. In this case the convergence is uniform with respect to $x \in [0,1]$, and hence one has*

$$\lim_{n\to\infty} \left\|\left(WN_n f\right) - f\right\|_{C[0,1]} = 0.$$

Now, let us consider the following Peetre's $K-$functional:

(7) $$K_2(f,\delta) := \inf_{g\in W^2} \{\left\|f-g\right\|_{C[0,1]} + \delta\left\|g''\right\|_{C[0,1]}\},$$

where $\delta > 0$ and $W^2 = \{g \in C[0,1] : g', g'' \in C[0,1]\}$. Then there exists an absolute constant $C > 0$ such that

(8) $$C^{-1}\omega_2(f,\sqrt{\delta}) \leq K_2(f,\delta) \leq C\omega_2(f,\sqrt{\delta}),$$

where

(9) $$\omega_2(f,\sqrt{\delta}) := \sup_{0<h\leq\sqrt{\delta}} \sup_{x\in[0,1]} \mid f(x+2h) - 2f(x+h) + f(x) \mid$$

is the second order modulus of smoothness of f. (see [8])

Theorem 3.5. *Let $f \in C[0,1]$ and let $\psi \in L_\infty(\mathbb{R})$ be a father wavelet satisfies **a)-c)**. Then*

$$\lim_{n\to\infty} \left(WN_n f\right)(x) = f(x),$$

and

$$\left|\left(WN_n f\right)(x) - f(x)\right| \leq (K+1)\, K_2\left(f; \frac{M_2(b_R) + \lambda^2 O_R + 2\lambda M_1(b_R)}{n^2}\right),$$

where $K = \lambda\left\|\psi\right\|_\infty$ and $K_2(f;\delta)$ is the Peetre's K-functional.

Proof. Let $g \in W^2$. By Taylor's theorem, we have

$$g(t) = g(x) + g'(x)(t-x) + \int_x^t (t-v)g''(v)dv, \;\; t \in [0,1].$$

In view of Remark 2 and (6), applying WN_n to the both sides of the above equation, we have

$$\begin{aligned}
\left|(WN_n g)(x) - g(x)\right| &= \left|\left(WN_n\left(g'(x)(t-x) + \int_x^t (t-v)g''(v)dv\right)\right)(x) - g(x)\right| \\
&\leq \frac{\sum\limits_{k=0}^{n} b_R(nx-k) \int\limits_{\mathbb{R}} \left|\int_x^{\frac{u+k}{n}} \left(\frac{u+k}{n} - v\right) g''(v)dv\right| |\psi(u)|\, du}{\sum\limits_{k=0}^{n} b_R(nx-k)} \\
&\leq \frac{\sum\limits_{k=0}^{n} b_R(nx-k) \int\limits_{0}^{\lambda} \left[\int_x^{\frac{u+k}{n}} \left|\frac{u+k}{n} - v\right| |g''(v)|\, dv\right] |\psi(u)|\, du}{\sum\limits_{k=0}^{n} b_R(nx-k)} \\
&\leq \lambda \|\psi\|_\infty \|g''\|_{C[0,1]} \frac{\sum\limits_{k=0}^{n} b_R(nx-k)\left(\frac{\lambda+k}{n} - x\right)^2}{\sum\limits_{k=0}^{n} b_R(nx-k)} \\
&= \lambda \|\psi\|_\infty \|g''\|_{C[0,1]} \frac{\sum\limits_{k=0}^{n} b_R(nx-k)\left[\left(\frac{k}{n} - x\right)^2 + \frac{\lambda^2}{n^2} + 2\frac{\lambda}{n}\left(\frac{k}{n} - x\right)\right]}{\sum\limits_{k=0}^{n} b_R(nx-k)} \\
&\leq \lambda \|\psi\|_\infty \|g''\|_{C[0,1]} \left[\frac{M_2(b_R)}{n^2} + \frac{\lambda^2}{n^2} O_R + 2\frac{\lambda}{n}\frac{M_1(b_R)}{n}\right] \\
&= \frac{\lambda \|\psi\|_\infty \|g''\|_{C[0,1]}}{n^2}\left[M_2(b_R) + \lambda^2 O_R + 2\lambda M_1(b_R)\right].
\end{aligned}$$

Hence, taking infimum on the right hand side over all $g \in W^2$ and using (7), we get

$$\begin{aligned}
\left|(WN_n f)(x) - f(x)\right| &\leq \inf_{g\in W^2}\{\|WN_n(f-g)\|_\infty + \|f-g\|_\infty + |(WN_n g)(x) - g(x)|\} \\
&\leq \inf_{g\in W^2}\{(\lambda\|\psi\|_\infty + 1)\|f-g\|_\infty + \frac{\lambda\|\psi\|_\infty\left[M_2(b_R) + \lambda^2 O_R + 2\lambda M_1(b_R)\right]}{n^2}\|g''\|_\infty\} \\
&\leq (K+1)\inf_{g\in W^2}\{\|f-g\|_\infty + \frac{M_2(b_R) + \lambda^2 O_R + 2\lambda M_1(b_R)}{n^2}\|g''\|_\infty\} \\
&= (K+1) K_2\left(f; \frac{M_2(b_R) + \lambda^2 O_R + 2\lambda M_1(b_R)}{n^2}\right),
\end{aligned}$$

here $K = \lambda \|\psi\|_\infty$. □

Theorem 3.6. *Let $f \in C[0,1]$, $\psi \in L_\infty(\mathbb{R})$ be a father wavelet satisfies* ***a)-c)*** *and $\alpha \in (0,2)$ be fixed real number. Then*

$$\omega_2(f;t) = \mathcal{O}(t^\alpha) \Rightarrow |(WN_n f)(x) - f(x)| = \mathcal{O}(1/n)^\alpha$$

holds true.

Proof. In view of (7), (9) and the relation (8) betwen modulus of smoothess and Peetre's K-functional, we have from Theorem 4

$$\begin{aligned}
\left|(WN_n f)(x) - f(x)\right| &\leq (K+1) K_2\left(f; \frac{M_2(b_R) + \lambda^2 O_R + 2\lambda M_1(b_R)}{n^2}\right) \\
&\leq (K+1) C\omega_2\left(f; \sqrt{\frac{M_2(b_R) + \lambda^2 O_R + 2\lambda M_1(b_R)}{n^2}}\right) \\
&\leq (K+1) C\left(\frac{M_2(b_R) + \lambda^2 O_R + 2\lambda M_1(b_R)}{n^2}\right)^{\alpha/2}.
\end{aligned}$$

□

4. Graphical representations

Now, we will give some graphical examples for these approach, namely convergence to functions by means of wavelet type Neural Network operators $(WN_n f)(x)$.

We note that in all the following Figures, the graph with the red line belongs to the target function.

Example 4.1. *Let $f(x) = x^2$, and take the activation function as a Ramp function for the neural network operators. We consider a special case of the wavelet type Neural Network operators $(WN_n f)(x)$, namely Kantorovich type Neural Network operators. Then one has for $n = 3, 5$ and for $n = 20$.*

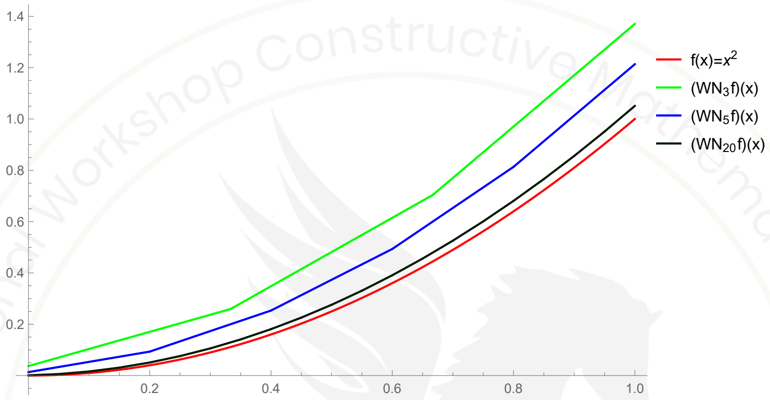

Figure 1. Approximation of $f(x) = x^2$ by Kantorovich type NN operator activated by Ramp function, for $n = 3, 5$ and $n = 20$.

Example 4.2. *Let $f(x) = x^2$, and take the activation function as a Ramp function for the neural network operators. We consider the wavelet type Neural Network operators $(WN_n f)(x)$ constructed by using Haar scaling function. Then one has for $n = 3, 5$ and for $n = 20$.*

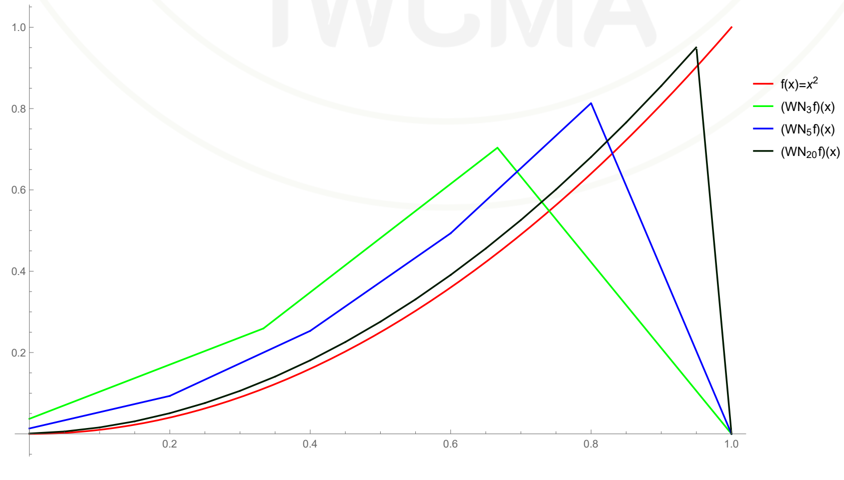

Figure 2. Approximation of $f(x) = x^2$ by Haar Wavelet type NN operator activated by Ramp function, for $n = 3, 5$ and $n = 20$.

Example 4.3. *Let $f(x) = x^2$, and take the activation function as a Ramp function for the neural network operators. We consider the wavelet type Neural Network operators $(WN_n f)(x)$ constructed by using Shannon wavelet function. Then one has for $n = 15, 36$ and for $n = 55$.*

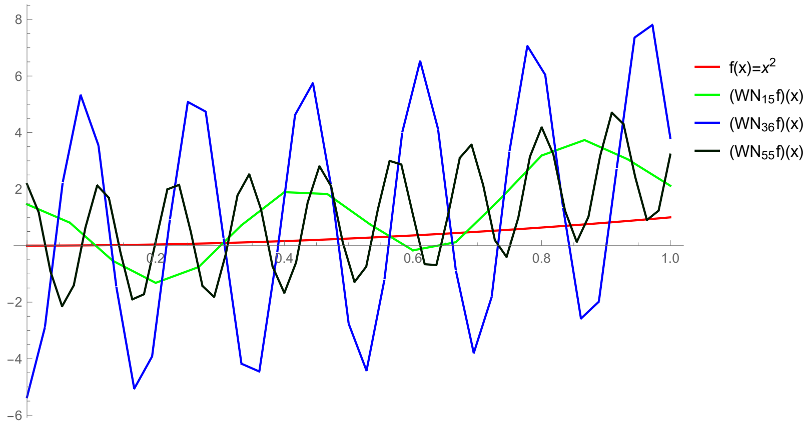

FIGURE 3. Approximation of $f(x) = x^2$ by Shannon Wavelet type NN operator activated by Ramp function, for $n = 15, 36$ and $n = 55$.

Example 4.4. *Let $f(x) = x - x^2$, and take the activation function as a Ramp function for the neural network operators. We consider a special case of the wavelet type Neural Network operators $(WN_n f)(x)$, namely Kantorovich type Neural Network operators. Then one has for $n = 3, 5$ and for $n = 20$.*

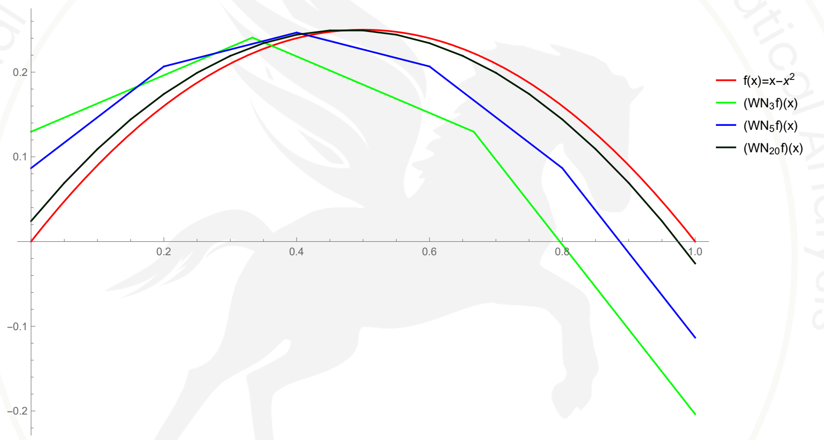

FIGURE 4. Approximation of $f(x) = x - x^2$ by Kantorovich type NN operator activated by Ramp function, for $n = 3, 5$ and $n = 20$.

Example 4.5. *Let $f(x) = x - x^2$, and take the activation function as a Ramp function for the neural network operators. We consider the wavelet type Neural Network operators $(WN_n f)(x)$ constructed by using Haar scaling function. Then one has for $n = 3, 6$ and for $n = 15$.*

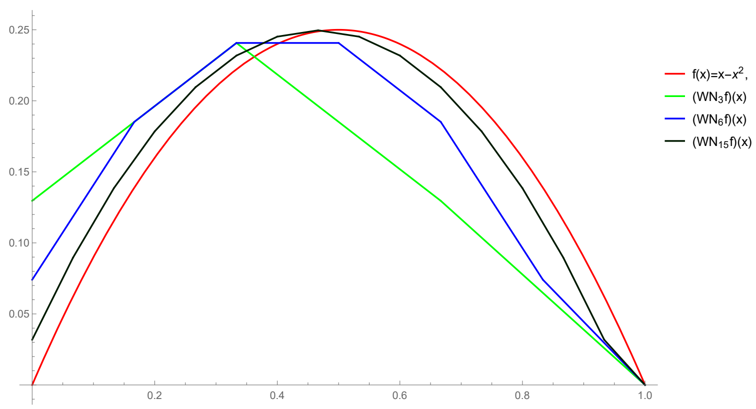

FIGURE 5. Approximation of $f(x) = x - x^2$ by Haar Wavelet type NN operator activated by Ramp function, for $n = 3, 6$ and $n = 15$.

Example 4.6. *Let $f(x) = x - x^2$, and take the activation function as a Ramp function for the neural network operators. We consider the wavelet type Neural Network operators $(WN_n f)(x)$ constructed by using Shannon wavelet function. Then one has for $n = 30, 60$ and for $n = 150$.*

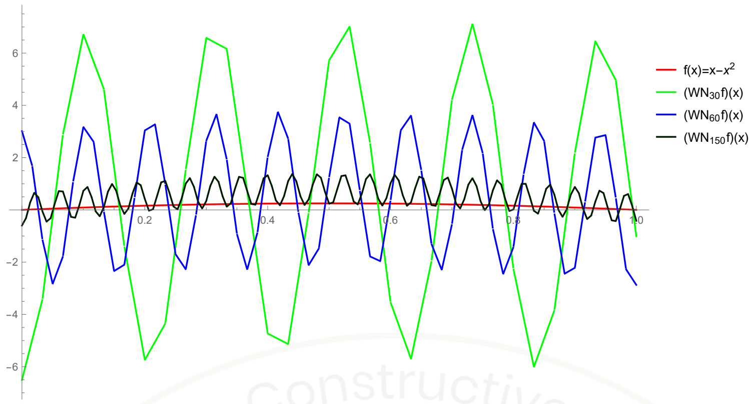

FIGURE 6. Approximation of $f(x) = x - x^2$ by Shannon Wavelet type NN operator activated by Ramp function, for $n = 30, 60$ and $n = 150$.

Note 2. Since the compactly supported Daubechies wavelets are also an unconditional orthonormal base of $L^p(\mathbb{R})$, this allows us to investigate the convergence problem on $L^p(\mathbb{R})$ by means of our wavelet type NN operators (5). These will be the future studies and investigations on this topic.

Moreover, neurocomputing processes deal with multidimensional data,.then clearly multivariate Neural Networks have special interest.

Therefore, it is planned as further studies to expand the existing theory to cover the case of multivariate functions and also to make image processing applications.

References

[1] Agratini, O.: Construction of Baskakov-type operators by wavelets. Rev. Anal. Numér. Théor. Approx. 26 (1997), no. 1-2, 3–11.
[2] Anastassiou, G. A. : Rate of convergence of some neural network operators to the unit-univariate case. Journal of Mathematical Analysis and Applications, 212, 237–262, (1997).
[3] Anastassiou, G. A.: Univariate sigmoidal neural network approximation. Journal of Computational Analysis and Applications, 14(4), 659-690, (2012).
[4] Anastassiou, G.A.: Univariate hyperbolic tangent neural network approximation. Math. Comput. Model. 53, 1111–1132 (2011).
[5] Anastassiou, G.A.: Intelligent Systems: Approximation by Artificial Neural Networks. Intelligent Systems Reference Library, vol. 19. Springer, Berlin (2011).
[6] Anastassiou, G.A., Coroianu, L., Gal, S.G.: Approximation by a nonlinear Cardaliaguet–Euvrard neural network operator of max-product kind. J. Comput. Anal. Appl. 12(2), 396–406 (2010).
[7] Bajpeji, S., Kumar, A.S.: Approximation by exponential sampling type neural network operators, Analysis and Mathematical Physics (2021) 11:108.
[8] Butzer, P. L. and Nessel, R. J.: Fourier Analysis and Approximation, V.1, Academic Press, New York, London, 1971.
[9] Cantarini, M., Costarelli, D., Vinti, G.: Asymptotic Expansion for Neural Network Operators of the Kantorovich Type and High Order of Approximation, Mediterr. J. Math. (2021) 18:66.
[10] Cao, F., Chen, Z.: Scattered data approximation by neural networks operators. Neurocomputing 190, 237–242 (2016).
[11] Cao, F., Xu, Z.: Simultaneous Lp-approximation order for neural networks. Neural Netw. 18(7), 914–923 (2005).
[12] Cardaliaguet, P., Euvrard, G.: Approximation of a function and its derivative with a neural network. Neural Netw. 5(2), 207–220 (1992).
[13] Chen, Z., Cao, F.: The construction and approximation of a class of neural networks operators with ramp functions. J. Comput. Anal. Appl. 14(1), 101–112 (2012).
[14] Costarelli, D., Spigler, R.: Convergence of a family of neural network operators of the Kantorovich type. J. Approx. Theory 185, 80–90 (2014).
[15] Costarelli, D., Spigler, R.: Approximation by series of sigmoidal functions with applications to neural networks. Annali di Matematica (2015) 194:289–306.
[16] Costarelli, D., Spigler, R., Vinti, G.: A survey on approximation by means of neural network operators. J. Neurotechnol. 1(1), 29–52 (2016).

[17] Costarelli, D., Vinti, G.: Approximation by max-product neural network operators of Kantorovich type, Results Math. 69(3) (2016), 505–519.

[18] Costarelli, D., Vinti, G.: Max-product neural network and quasi-interpolation operators activated by sigmoidal functions, J. Approx. Theory 209 (2016), 1–22.

[19] Costarelli, D., Vinti, G.: Convergence for a family of neural network operators in Orlicz spaces. Math. Nachr. 290(2–3), 226–235 (2017).

[20] Costarelli, D., Vinti, G.: Pointwise and uniform approximation by multivariate neural network operators of the max-product type. Neural Netw. 81, 81–90 (2016).

[21] Costarelli, D., Vinti, G.: Voronovskaja formulas for high order convergence neural network operator with sigmoidal functions. Mediterr. J. Math. 17 (article numb. 77), https://doi.org/10.1007/s00009-020-01513-7 (2020).

[22] Cucker, F., Zhou, D.X.: Learning Theory an Approximation Theory Viewpoint. Cambridge University Press, Cambridge (2007).

[23] Cybenko, G.: Approximation by superpositions of a sigmoidal function. Mathematics of Control, Signals, and Systems, 2, 303–314, (1989).

[24] Daubechies I.: Orthonormal bases of compactly supported wavelets, Comm. Pure Appl. Math. 41 (1988), 909-996.

[25] Daubechies I.: Ten Lectures on Wavelets, CBMS-NSF Series in Appl. Math. 61, SIAM Publ. Philadelphia, 1992.

[26] Gonska H. H. and Zhou D. X.: Using wavelets for Szász-type operators., Rev. Anal. Numér. Théor. Approx. 24 (1995), no. 1-2, 131–145.

[27] Karsli, H.: Extension of the generalized Bézier operators by wavelet, General Math., Vol. 30, No. 2 (2022), 3–15.

[28] Karsli H.: On Wavelet Type Bernstein Operators. Carpathian Math. Publ. 2023, 15 (1), 212-221.

[29] Karsli, H.: On wavelet type generalized Bézier Operators, Math. Found. Comp. 2023, 6(3): 439-452.

[30] Karsli, H.: On wavelet type generalized sampling operators, submitted.

[31] Kelly S., Kon M, and Raphael L.: Pointwise convergence of wavelet expansions, Bull. Amer. Math. Soc., Vol. 30 (1994), pp. 87-94.

[32] Łenski W., Szal B.: Approximation of Integrable Functions by Wavelet Expansions. Results Math 72, 1203–1211 (2017).

[33] Smale, S., Zhou, D.X.: Learning theory estimates via integral operators and their approximations. Construct. Approx. 26(2), 153–172 (2007).

[34] Unser, M.: Ten good reasons for using spline wavelets. Wavelets Appl. Signal Image Process. 3169(5), 422–431 (1997).

[35] Walter G.G.: Pointwise convergence of wavelet expansions. J. Approx. Theory 80(1), 108-118 (1995).

[36] Verhulst, P.F.: Notice Sur la Loi Que la Population Poursuit dans Son Accroissement, Correspondance Mathematique et Physique 10, 113-121 (1838).

[37] Verhulst, P.F.: Recherches Mathematiques sur la Loi D'accroissement de la Population (Mathematical Researches into the Law of Population Growth Increase), Nouveaux Memoires de l'Academie Royale des Sciences et Belles-Lettres de Bruxelles 18, 1-42 (1845).

Harun Karsli

Bolu Abant Izzet Baysal University, Faculty of Science and Arts, Department of Mathematics, 14030, Gölköy-Bolu, Turkey

Email(s): karsli_h@ibu.edu.tr

Comparison of imputation and weighting methods in estimation of a finite population mean under random non-response

Nelson Kiprono Bii*[1] and Christopher Ouma Onyango[2]

Abstract. Developing estimators of finite population parameters such as mean, variance and asymptotic mean squared error has been one of the core objectives of sample survey theory and practice. Sample survey practitioners need to assess the properties of these estimators so that better ones can be adopted. In survey sampling, the occurrence of non-response affects inference and optimality of the estimators of finite population parameters. It introduces bias and may cause samples not to follow the the distributions determined by the original sampling design. To compensate for random non-response, imputation methods and weighting techniques can be used. In this paper, a comparison between these two methods of compensating for non-response has been done in two-stage cluster sampling. Simulation results reveal tighter confidence interval lengths, smaller mean squared error values for the estimators developed under the weighting method than its rival estimators obtained using imputation method. Under mild assumptions, the weighting method is shown to be more efficient than the imputation techniques in estimating a finite population mean.

2020 Mathematics Subject Classifications: 62D05, 62D10, 62G05

Keywords: Sampling With Replacement, Two-Stage Cluster Sampling, Bias and Mean Squared Error.

1. Introduction

In the estimation of a finite population mean, a lot of significance is attached to efficient and cost-effective survey sampling designs in sample surveys, see for instance, (*cf.* [1]) . Careful design of samples based on random selection with known probabilities of population elements should be considered. This gives a target sample of intended respondents where each may provide responses to a set of survey questions that results in an array of responses (*cf.* [23]) observed that non-response occurs if some of the expected responses are missing, for instance where a whole vector of responses is missing for some sampled units or where responses are obtained for some questions and not to others in the sample selected.
The basis for statistical inference is therefore formed by a sampling design that provides a link between a sample and the population. As observed by (*cf.* [10]), a good sample survey practise and efficient methods of compensating for non-response should be adopted.

1.1. Preliminaries.

In sample surveys, non-response introduces bias in the estimation of a finite population mean. It also causes samples to fail to follow the distributions determined by the original sampling design. The use of regression models is recognized as one of the procedures for reducing bias due to non-response using auxiliary information, for details see (*cf.* [2]). In practise, information on the variables of interest is not available for non-respondents but information on auxiliary variables may be available for non-respondents. To reduce the bias and variance due to non-response, (*cf.* [13]) noted that it is desirable to incorporate auxiliary data into the estimation process where the response probabilities are mostly taken to be correlated with certain characteristics such as age, race and income for a human population survey.

1.2. Methods of Compensating for Non-response.

Imputation techniques and weighting method are discussed in this paper in the following subsections.

1.2.1. *Imputation Methods.* Imputation entails compensating for non-response values by proxy values as observed by (*cf.* [20]). Imputation techniques have been used to account for non-response in the study variable in the estimation of finite population mean. For instance, (*cf.* [21]) applied compromised method of imputation to estimate a finite population mean under two stage cluster sampling, though the method produced a large bias.

Following the procedure by (*cf.* [9]), let $\overline{Y} = \frac{1}{N}\sum_1^N Y_i$ represents the finite population mean to be estimated. Besides, let a simple random sample, say, S with replacement be drawn from the population, $\theta = 1, 2, \ldots, N$ to estimate Y. The sample S of units has r responding units $(r < n)$ making a set R and $(n-r)$ non-responding units with the subspace $(n-r)$ having the symbol R^C in the population. For every unit $i \in R$, the value of the survey variable y_i is obtained.

However, imputed values are to be derived for the non-response set of units, that is, for every $i \in R^C$, since the y_i values are missing. It is assumed that auxiliary data x_i are known for every $i \in S$. The value of the auxiliary, x_i, imputes the non-response values when $i \in R^C$, that is, for a sample S assume that the data $x_s = x_i : i \in s$ are given and $S = R \cup R^C$. Using this set up, ratio method of imputation, an example of different imputation techniques, can be defined as in equation (2). Using the notations of (*cf.* [17]), if the i^{th} unit is to be imputed, the value $\hat{b}x_i$ is obtained, where $\hat{b}x_i = \frac{\sum_{i\in R} y_i}{\sum_{i\in R} x_i}$. The data after imputation becomes

$$y_{\cdot i} = \begin{cases} y_i, \text{ if } i \in R \\ \hat{b}x_i, \; i \in R^C \end{cases} \tag{1}$$

For details, see (*cf.* [21]). The imputation ratio estimator is given by

$$\overline{y}_{RAT} = \overline{y}_r \frac{\overline{x}_n}{\overline{x}_r} \tag{2}$$

where $\overline{x}_n = \frac{1}{n}\sum_{i\in s} x_i$, $\overline{x}_r = \frac{1}{r}\sum_{i\in R} x_i$ and $\overline{y}_r = \frac{1}{r}\sum_{i\in R} y_i$. The bias and the mean squared error of the imputation ratio estimator due to (*cf.* [21]) are given below.

$$B(\overline{y}_{RAT}) = (\frac{1}{r} - \frac{1}{n})\overline{Y}\Big(C_x^2 - \rho C_x C_y\Big) \tag{3}$$

where $C_x = \frac{s_x}{\overline{X}}$, $C_y = \frac{s_y}{\overline{Y}}$ and $\rho = \frac{S_{xy}}{S_x S_y}$. S_x and S_y are the standard deviations of X and Y values respectively while S_{xy} is the co-variance between X and Y. The MSE is given by

$$MSE(\overline{y}_{RAT}) = (\frac{1}{n} - \frac{1}{N})S_y^2 + \Big(\frac{1}{r} - \frac{1}{n}\Big)\Big[S_y^2 + R_1^2 S_x^2 - 2R_1 S_{xy}\Big] \tag{4}$$

where $R_1 = \frac{\overline{Y}}{\overline{X}}$. (*cf.* [1]) observed that though this method of imputation is better than other existing techniques like mean and compromised method of imputation, its bias and MSE are still large compared to rival approaches of compensating for non-response such as weighting techniques.

1.2.2. *Weighting Method.* It has been observed by authors such as (*cf.* [16]) that non-response causes loss of observations and therefore weighting means that the weights are increased for all or almost all of the elements that fail to respond in a survey. For instance, (*cf.* [4]) and (*cf.* [11]) discussed a modified Horvitz-Thompson estimator to correct for non-response problem using the weighting strategy. The estimator used was defined by

$$\hat{\hat{y}}_{HT} = N^{-1}\sum_{k=1}^{N}\big(\phi_k \hat{\hat{p}}_{HT}\big)^{-1} y_k \tau_k \tag{5}$$

where $\hat{\hat{p}}_{HT}$ is given by

$$\hat{\hat{p}}_{HT} = N^{-1}\sum_{k=1}^{N}\phi_k^{-1}\gamma_k$$

where $y_k, k \in U$ is the value of the k^{th} survey variable taken from a sample s selected from a finite population, $U = \big(1, 2, \cdots, N\big)$, ϕ_k is the inclusion probability given by $\phi_k = p_r(k \in s)$, γ_k is the value of the k^{th} respondent in the sample selected, s. The estimator $\hat{\hat{y}}_{HT}$ adjusts the weights by an unbiased estimator $\hat{\bar{p}}_{HT}$, of the response probabilities of the population mean which is given by

$$\bar{P}_N = N^{-1}\sum_{i=1}^{N} p_i \tag{6}$$

Thus an approximate bias of the estimator $\hat{\bar{y}}_{HT}$ is given by

$$Bias(\hat{\bar{y}}_{HT}) = E(\hat{\bar{y}}_{HT}) - \bar{Y} = \bar{Y}^{*} - \bar{Y} = \hat{\bar{P}}_N^{-1} C_p Y \tag{7}$$

where

$$\bar{Y}^{*} = N^{-1}P_N^{-1}\sum_{i=1}^{N} p_i y_i \tag{8}$$

and

$$C_p Y = N^{-1}\sum_{i=1}^{N}\big(p_i - \bar{P}_N\big)\big(y_i - \bar{Y}\big). \tag{9}$$

If C_pY is close to zero, the bias will be small, for more details see (*cf.* [20]). The adjusted Horvitz-Thompson estimator, $\hat{\bar{y}}_{HT}$, is an illustration of re-weighting measurements of respondents without using auxiliary information. However, in this paper, auxiliary information is used in the estimation procedure.

The population mean, $\overline{Y} = \frac{1}{N}\sum_1^N Y_i$, is estimated by selecting a sample of size n at random with replacement. If the responding units of item y are independent so that the probability of unit j responding in cluster i is $p_{ij}(i = 1, 2, \cdots, n; j = 1, 2, \cdots, m)$, then following the work of (*cf.* [18]), a weighted estimator, $\hat{\bar{y}}_I$, for $\bar{Y}$ is given by

$$\hat{\bar{y}}_I = \frac{1}{\sum_{i,j\in s} w_{ij}}\left[\sum_{i,j\in s_r} w_{ij}y_{ij} + \sum_{i,j\in s_m} w_{ij}y_{ij}^{*}\right] \tag{10}$$

where $w_{ij} = \frac{1}{\pi_{ij}}$ gives the survey weight tied to unit j in cluster i and $\pi_{ij} = p[i, j \in s]$ is its probability of inclusion, s_r, is the set of r responding units to item y, s_m is the set of m units that failed to respond to item y so that $r + m = n$ while y_{ij}^{*} is the value imputed so that the missing value y_{ij} is compensated for, (*cf.* [2]).

2. Main results

2.1. The proposed Estimator of Finite Population Mean. Consider a finite population of size N consisting of M clusters with N_j elements in the j^{th} cluster. A sample of m clusters is selected so that n_{1i} units respond and n_{2i} units fail to respond. Let y_{ij} denote the value of the survey variable y for unit j in cluster i, for $i = 1, 2, \cdots, N, j = 1, 2, \cdots, N_i$ and let population mean be given by

$$\overline{\overline{Y}} = \frac{1}{N}\sum_{i=1}^{M}\sum_{j=1}^{N_i} Y_{ij}\,. \tag{11}$$

The proposed estimator is given by

$$\hat{\overline{\overline{Y}}}_{INW} = \frac{1}{M}\Big\{\frac{1}{n_{1i}}\sum_{i\in s}\sum_{j\in s} y_{ij} + \frac{1}{n_{2i}}\sum_{i\in s}\sum_{j\notin s}\hat{y}_{ij}\Big\}. \tag{12}$$

where $\hat{y}_{ij}$ is an estimator of the non-response component of the sample. Assuming auxiliary information, X_{ij}, is known throughout, $\hat{y}_{ij}$ can be obtained using the improved Nadaraya-Watson regression technique by

$$\hat{y}_{ij} = m_{INW}(\hat{x}_{ij}) = \frac{\sum_{i\in s}\sum_{j\in s}\frac{1}{\lambda_{ij}}K\left(\frac{x - X_{ij}}{\lambda_{ij}b}\right)Y_{ij}}{\sum_{i\in s}\sum_{j\in s}\frac{1}{\lambda_{ij}}K\left(\frac{x - X_{ij}}{\lambda_{ij}b}\right)} \tag{13}$$

so that the estimator of finite population mean can be re-written as

$$\hat{\bar{\bar{Y}}}_{INW} = \frac{1}{M}\Big\{\frac{1}{n_{1i}}\sum_{i\in s}\sum_{j\in s} y_{ij} + \frac{1}{n_{2i}}\sum_{i\in s}\sum_{j\notin s} m_{INW}(\hat{x}_{ij})\Big\}. \tag{14}$$

A special case where $n_{1i} = n_{2i} = n$ is assumed in this study. This simplifies mathematical computations so that equation (15) can be re-written as

$$\hat{\bar{\bar{Y}}}_{INW} = \frac{1}{Mn}\Big\{\sum_{i\in s}\sum_{j\in s} y_{ij} + \sum_{i\in s}\sum_{j\notin s} m_{INW}(\hat{x}_{ij})\Big\}, \tag{15}$$

where $m_{INW}(\hat{x}_{ij})$ is the improved Nadaraya-Watson kernel regression estimator given in (**??**), which is a weighted sum of the values of the survey variable Y_{ij}'s. Data is generated using a regression model given by

$$\hat{Y}_{ij} = m(\hat{x}_{ij}) + \hat{e}_{ij} \tag{16}$$

where $m(.)$ is an unknown smooth function of auxiliary random variables, X_{ij}. It is assumed that the error term, $\hat{e}_{ij}$, satisfies the following conditions:

$$E(\hat{e}_{ij}) = 0, Var(\hat{e}_{ij}) = \sigma^2_{ij}, Cov(\hat{e}_i, \hat{e}_j) = 0, for \quad i \neq j \tag{17}$$

Hence the unspecified function of the auxiliary random variables, $m(\hat{x}_{ij})$, is replaced by the improved Nadaraya-Watson kernel estimator, $m_{INW}(\hat{x}_{ij})$. The estimator can be re-written as

$$m_{INW}(\hat{x}_{ij}) = \sum_{i\in s}\sum_{j\in s} w(x_{ij})Y_{ij}. \tag{18}$$

where $w(x_{ij}) = \dfrac{\frac{1}{\lambda_{ij}}K\left(\frac{x-X_{ij}}{\lambda_{ij}b}\right)}{\sum_{i\in s}\sum_{j\in s}\frac{1}{\lambda_{ij}}K\left(\frac{x-X_{ij}}{\lambda_{ij}b}\right)}$ are the improved Nadaraya-Watson kernel weights where $K(.)$ is a given kernel function assumed to be symmetrical; b is a smoothing parameter while λ_{ij} is the local bandwidth given by

$$\lambda_{ij} = \{m(X_{ij}/a)\}^{-\alpha}. \tag{19}$$

where α is a sensitivity parameter which satisfies $0 \leq \alpha \leq 1$. It has been suggested by (*cf.* [12]) that taking $\alpha = \frac{1}{2}$ produce good results. Since the choice of the kernel function is not critical for the performance of the kernel regression estimator, a simplified Gaussian kernel with mean 0 and variance 1 is used in this study. This is given by

$$K(w) = \frac{1}{\sqrt{2\pi}}e^{-\left(\frac{w^2}{2}\right)} = \frac{1}{\sqrt{2\pi}}e^{-\left(\frac{\left(\frac{x-X_{ij}}{\lambda_{ij}b}\right)^2}{2}\right)} \tag{20}$$

In this case, the improved Nadaraya-Watson kernel estimation at any point x_{ij} is given by

$$\hat{y}_{ij} = m_{INW}(\hat{x}_{ij}) = \frac{\sum_i\sum_j \frac{1}{\sqrt{2\pi}}e^{-\left(\frac{\left(\frac{x-X_{ij}}{\lambda_{ij}b}\right)^2}{2}\right)}Y_{ij}}{\sum_i\sum_j \frac{1}{\sqrt{2\pi}}e^{-\left(\frac{\left(\frac{x-X_{ij}}{\lambda_{ij}b}\right)^2}{2}\right)}} \tag{21}$$

where b is the bandwidth while λ_{ij} is given in equation (**??**) due to (*cf.* [7]).
This provides a way of estimating the non-response values of the survey variable Y_{ij}, in the i^{th} cluster given the auxiliary values x_{ij}, for a specified kernel function.

2.2. **The Asymptotic Bias of the Proposed Estimator.** The expected value of the proposed estimator is given by

$$E(\hat{\bar{\bar{Y}}}_{INW}) = \frac{1}{Mn}\left\{\sum_{i=1}^{n}\sum_{j=1}^{m} Y_{ij} + \sum_{i=n+1}^{N}\sum_{j=m+1}^{M} m_{INW}(\hat{x}_{ij})\right\} \tag{22}$$

Re-writing equation (13) using the property of symmetry associated with Nadaraya-Watson estimator,

$$m_{INW}(\hat{x}_{ij}) = \frac{\sum_{i\in s}\sum_{j\in s} K\left(\frac{X_{ij}-x_{ij}}{\lambda_{ij}b}\right) Y_{ij}}{\sum_{i\in s}\sum_{j\in s} K\left(\frac{X_{ij}-x_{ij}}{\lambda_{ij}b}\right)}, i = 1, 2, \cdots, n; j = 1, 2, \cdots, m \tag{23}$$

Following the procedure by (*cf.* [22]), equation (23) can be re-written as

$$m_{INW}(\hat{x}_{ij}) = \frac{1}{g(\hat{x}_{ij})}\left[\frac{1}{mn(\lambda_{ij}b)}\sum_i\sum_j K\left(\frac{X_{ij}-x_{ij}}{\lambda_{ij}b}\right) Y_{ij}\right] \tag{24}$$

where $g(\hat{x}_{ij})$ is the estimated marginal density of auxiliary variables X_{ij}. The bias of the estimator can be written as

$$Bias\,(\hat{\overline{\overline{Y}}}_{INW}) = E(\hat{\overline{\overline{Y}}}_{INW} - \overline{\overline{Y}}) \tag{25}$$

$$\begin{aligned} Bias\,(\hat{\overline{\overline{Y}}}_{INW}) = E\Bigg\{&\frac{1}{Mn}\Bigg[\sum_{i=1}^{n}\sum_{j=1}^{m} Y_{ij} + \sum_{i=n+1}^{N}\sum_{j=m+1}^{M} m_{INW}(\hat{x}_{ij})\Bigg] \\ &- \frac{1}{Mn}\Bigg[\sum_{i=1}^{n}\sum_{j=1}^{m} Y_{ij} + \sum_{i=n+1}^{N}\sum_{j=m+1}^{M} Y_{ij}\Bigg]\Bigg\} \end{aligned} \tag{26}$$

which reduces to

$$Bias\,(\hat{\overline{\overline{Y}}}_{INW}) = \frac{1}{Mn}E\left\{\sum_{i=n+1}^{N}\sum_{j=m+1}^{M} m_{INW}(\hat{x}_{ij}) - \sum_{i=n+1}^{N}\sum_{j=m+1}^{M} Y_{ij}\right\} \tag{27}$$

Re-writing the regression model given by $Y_{ij} = m(X_{ij}) + e_{ij}$ as

$$Y_{ij} = m(x_{ij}) + \big[m(X_{ij}) - m(x_{ij})\big] + e_{ij} \tag{28}$$

and substituting it in equation (24) gives

$$\begin{aligned} m_{INW}(\hat{x}_{ij}) = \frac{1}{g(\hat{x}_{ij})}\Bigg[&\frac{1}{mn(\lambda_{ij}b)}\sum_i\sum_j K\Big(\frac{X_{ij}-x_{ij}}{\lambda_{ij}b}\Big)\Big(m(x_{ij}) + \big[m(X_{ij}) - m(x_{ij})\big] \\ &+ e_{ij}\Big)\Bigg] \end{aligned} \tag{29}$$

Hence the first term in equation (27) before taking expectation is given as:

$$\begin{aligned} &\frac{1}{Mn}\left\{\frac{\frac{1}{mnb}\sum_{i=n+1}^{N}\sum_{j=m+1}^{M} K\left(\frac{X_{ij}-x_{ij}}{\lambda_{ij}b}\right) Y_{ij}}{g(\hat{x}_{ij})}\right\} \\ &\quad = \frac{1}{Mng(\hat{x}_{ij})}\left\{\frac{1}{mn(\lambda_{ij}b)}\sum_{i=n+1}^{N}\sum_{j=m+1}^{M} K\left(\frac{X_{ij}-x_{ij}}{\lambda_{ij}b}\right) m(x_{ij})\right. \\ &\quad + \frac{1}{mn(\lambda_{ij}b)}\sum_{i=n+1}^{N}\sum_{j=m+1}^{M} K\left(\frac{X_{ij}-x_{ij}}{\lambda_{ij}b}\right)\Big[m(X_{ij}) - m(x_{ij})\Big] \\ &\quad \left. + \frac{1}{mn(\lambda_{ij}b)}\sum_{i=n+1}^{N}\sum_{j=m+1}^{M} K\left(\frac{X_{ij}-x_{ij}}{\lambda_{ij}b}\right) e_{ij}\right\} \end{aligned} \tag{30}$$

Simplifying equation (30), the following is obtained:

$$\begin{aligned}&\frac{1}{Mn}\left\{\frac{1}{mn(\lambda_{ij}b)\,g(\hat{x}_{ij})}\sum_{i=n+1}^{N}\sum_{j=m+1}^{M}K\left(\frac{X_{ij}-x_{ij}}{\lambda_{ij}b}\right)Y_{ij}\right\}\\ &=\frac{1}{Mn}\left(\frac{1}{mn(\lambda_{ij}b)\,g(\hat{x}_{ij})}\right)\left\{\sum_{i=n+1}^{N}\sum_{j=m+1}^{M}g(\hat{x}_{ij})m(x_{ij})\right.\\ &\left.\quad+m_1(\hat{x}_{ij})+m_2(\hat{x}_{ij})\right\},\end{aligned}\tag{31}$$

where

$$m_1(\hat{x}_{ij})=\frac{1}{mn(\lambda_{ij}b)}\sum_{i=n+1}^{N}\sum_{j=m+1}^{M}K\left(\frac{X_{ij}-x_{ij}}{\lambda_{ij}b}\right)\Big[m(X_{ij})-m(x_{ij})\Big].\tag{32}$$

$$m_2(\hat{x}_{ij})=\frac{1}{mn(\lambda_{ij}b)}\sum_{i=n+1}^{N}\sum_{j=m+1}^{M}K\left(\frac{X_{ij}-x_{ij}}{\lambda_{ij}b}\right)e_{ij}.\tag{33}$$

Taking conditional expectation of equation (31) leads to

$$\begin{aligned}E\Big[\sum_{i=n+1}^{N}\sum_{j=m+1}^{M}m_{INW}(\hat{x}_{ij})/x_{ij}\Big]&=\frac{1}{Mn}E\Bigg[\frac{1}{mn(\lambda_{ij}b)}\sum_{i=n+1}^{N}\sum_{j=m+1}^{M}\Bigg[m(x_{ij})\\ &\quad+\frac{m_1(\hat{x}_{ij})}{g(\hat{x}_{ij})}+\frac{m_2(\hat{x}_{ij})}{g(\hat{x}_{ij})}\Bigg]\Bigg].\end{aligned}\tag{34}$$

The following theorem due to (*cf.* [8]) and applied by (*cf.* [19]) was used in obtaining asymptotic bias and variance of the estimator using conditional expectations.

Theorem 2.1. *Let $K(w)$ be a symmetric density function with $\int wk(w)dw=0$ and $\int w^2k(w)dw=k_2$. Assume n and N increase together such that $\frac{n}{N}\to\pi$ with $0<\pi<1$. Besides, assume the sampled and non-sampled values of x are in the interval $[c,d]$ and are obtained by densities d_s and d_{p-s} respectively where both are bounded away from zero on $[c,d]$ with continuous second derivatives. If for any variable $Z, E(Z/U=u)=A(u)+O(B)$ and $Var(Z/U=u)=O(C)$, then $Z=A(u)+O_p(B+C^{\frac{1}{2}})$.*

Using this theorem, the asymptotic bias can further be derived and simplified. From the conditions of the error term stated in (17), it follows that $E(e_{ij}/X_{ij})=0$. Therefore, $E\big[m_2(\hat{x}_{ij})=0\big]$. Thus, $E\big[m_1(\hat{x}_{ij})\big]$ can be obtained as follows:

$$\begin{aligned}E\sum_{i=n+1}^{N}\sum_{j=m+1}^{M}\Big[m_1(\hat{x}_{ij})\Big]&=\frac{1}{Mn}\Big(\frac{1}{mn(\lambda_{ij}b)}\Big)E\Bigg\{\sum_{i=n+1}^{N}\sum_{j=m+1}^{M}K\Bigg(\frac{X_{ij}-x_{ij}}{\lambda_{ij}b}\Bigg)\\ &\quad\times\Big[m(X_{ij})-m(x_{ij})\Big]\Bigg\}.\end{aligned}\tag{35}$$

Using substitution and change of variable technique given by

$$\left.\begin{array}{c} w=\frac{V-x_{ij}}{\lambda_{ij}b}\\ V=x_{ij}+(\lambda_{ij}b)w\\ dV=(\lambda_{ij}b)dw\end{array}\right\}\tag{36}$$

Equation (35) can be simplified to:

$$\begin{aligned}E\sum_{i=n+1}^{N}\sum_{j=m+1}^{M}\Big[m_1(\hat{x}_{ij})\Big]&=\frac{1}{Mn}\Bigg\{\frac{Mn-mn}{mn}\int k(w)\Big[m(x_{ij}+(\lambda_{ij}b)w)\\ &\quad-m(x_{ij})\Big]\int g\Big(x_{ij}+(\lambda_{ij}b)w\Big)dw\Bigg\}\end{aligned}\tag{37}$$

Using Taylor's series expansion about the point x_{ij}, the k^{th} order kernel can be derived as follows:

$$g\big(x_{ij} + (\lambda_{ij}b)w\big) = g(x_{ij}) + g\prime(x_{ij})(\lambda_{ij}b)w + \frac{1}{2}g\prime\prime(x_{ij})(\lambda_{ij}b)^2w^2 + \cdots + \frac{1}{k!}g^k(x_{ij})(\lambda_{ij}b)^kw^k + o((\lambda_{ij}b)^2). \tag{38}$$

Similarly,

$$m\big(x_{ij} + (\lambda_{ij}b)w\big) = m(x_{ij}) + m\prime(x_{ij})(\lambda_{ij}b)w + \frac{1}{2}m\prime\prime(x_{ij})(\lambda_{ij}b)^2w^2 + \cdots + \frac{1}{k!}m^k(x_{ij})(\lambda_{ij}b)^kw^k + o((\lambda_{ij}b)^2). \tag{39}$$

Therefore, expanding equation (37) up to order $o((\lambda_{ij}b)^2)$ and simplifying gives

$$\begin{aligned} E\sum_{i=n+1}^{N}\sum_{j=m+1}^{M}\big[m_1(\hat{x}_{ij})\big] = \frac{1}{Mn}\bigg\{&(\frac{Mn-mn}{mn})g(x_{ij})m\prime(x_{ij})(\lambda_{ij}b)\int wk(w)dw \\ &+(\frac{Mn-mn}{mn})g\prime(x_{ij})m\prime(x_{ij})(\lambda_{ij}b)^2\int w^2k(w)dw \\ &+(\frac{Mn-mn}{mn})\frac{1}{2}g(x_{ij})m\prime\prime(x_{ij})(\lambda_{ij}b)^2 \\ &\times\int w^2k(w)dw + o((\lambda_{ij}b)^2)\bigg\}. \end{aligned} \tag{40}$$

Using the conditions due to (*cf.* [8]) given by $\int_{-\infty}^{\infty} k(w)dw = 1$, $\int_{-\infty}^{\infty} wk(w)dw = 0$ and $\int_{-\infty}^{\infty} w^2k(w)dw = d_k$, the derivation in equation (40) can further be simplified to obtain:

$$\begin{aligned} E\sum_{i=n+1}^{N}\sum_{j=m+1}^{M}\Big[m_1(\hat{x}_{ij})\Big] = \frac{1}{Mn}(\frac{Mn-mn}{mn})\Big[&g\prime(x_{ij})m\prime(x_{ij}) \\ &+\frac{1}{2}g(x_{ij})m\prime\prime(x_{ij})\Big](\lambda_{ij}b)^2d_k + o((\lambda_{ij}b)^2). \end{aligned} \tag{41}$$

Hence the expected value of the second term in equation (34) then becomes:

$$\begin{aligned} E\sum_{i=n+1}^{N}\sum_{j=m+1}^{M}\big[m_1(\hat{x}_{ij})\big] = \frac{1}{Mn}\bigg\{(\frac{Mn-mn}{mn})\bigg[&\frac{1}{2g(\hat{x}_{ij})}m\prime\prime(x_{ij})g(x_{ij}) \\ &+\frac{g\prime(x_{ij})m\prime(x_{ij})}{g(\hat{x}_{ij})}\bigg](\lambda_{ij}b)^2d_k + o((\lambda_{ij}b)^2)\bigg\}. \end{aligned} \tag{42}$$

Simplifying equation (42) gives:

$$E\sum_{i=n+1}^{N}\sum_{j=m+1}^{M}\big[m_1(\hat{x}_{ij})\big] = \frac{1}{Mn}\bigg\{(\frac{Mn-mn}{mn})(\lambda_{ij}b)^2d_kC(x) + o((\lambda_{ij}b)^2)\bigg\}, \tag{43}$$

where $C(x) = [g(\hat{x}_{ij})]^{-1}\Big[\frac{1}{2}m\prime\prime(x_{ij})g(x_{ij}) + g\prime(x_{ij})m\prime(x_{ij})\Big]$ and $d_k = \int w^2k(w)dw$.
Using equation of the bias given in (25) and the conditional expectation in equation (34), the following equation for the conditional bias of the estimator was obtained:

$$Bias\,(\hat{\bar{\bar{Y}}}_{INW}/x_{ij}) = \frac{1}{Mn}\bigg\{(\frac{Mn-mn}{mn})(\lambda_{ij}b)^2d_kC(x) + o((\lambda_{ij}b)^2)\bigg\}. \tag{44}$$

In the next subsection, the asymptotic variance of the estimator is also derived.

2.3. **Asymptotic Variance of the Proposed Estimator.** Using equation (15), the conditional variance of the estimator is given as

$$Var\big(\hat{\bar{\bar{Y}}}_{INW}/x_{ij}\big) = Var\bigg\{\frac{1}{Mn}\bigg\{\sum_{i=1}^{n}\sum_{j=1}^{m}Y_{ij} + \sum_{i=n+1}^{N}\sum_{j=m+1}^{M}m_{INW}\,(\hat{x}_{ij})\bigg\}\bigg\} \tag{45}$$

$$= (\frac{1}{Mn})^2Var\bigg\{\sum_{i=1}^{n}\sum_{j=1}^{m}Y_{ij} + \sum_{i=n+1}^{N}\sum_{j=m+1}^{M}m_{INW}\,(\hat{x}_{ij})\bigg\}, \tag{46}$$

where $m_{INW}(\hat{x}_{ij})$ is given by

$$m_{INW}\Big(\hat{x}_{ij}\Big) = \frac{1}{g(\hat{x}_{ij})}\Bigg[\frac{1}{mn(\lambda_{ij}b)}\sum_i\sum_j K\Big(\frac{X_{ij}-x_{ij}}{\lambda_{ij}b}\Big)Y_{ij}\Bigg], \tag{47}$$

where $g(\hat{x}_{ij}) = \frac{1}{mn(\lambda_{ij}b)}\sum_i\sum_j K\left(\frac{x-X_{ij}}{\lambda_{ij}b}\right)$ is the estimated marginal density of auxiliary variables X_{ij}, for details see (*cf.* [15]). Re-writing the regression model $Y_{ij} = m(X_{ij}) + e_{ij}$ as $Y_{ij} = m(x_{ij}) + \Big[m(X_{ij}) - m(x_{ij})\Big] + e_{ij}$ and substituting in equation (47) leads to

(48)

$$Var\big(\hat{\overline{\overline{Y}}}_{INW}/x_{ij}\big) = \big(\frac{1}{Mn}\big)^2 Var\Bigg\{\sum_{i=1}^{n}\sum_{j=1}^{m}Y_{ij} + \Bigg(\frac{1}{mn(\lambda_{ij}b)\,g(\hat{x}_{ij})}\Bigg)\Bigg\{\sum_{i=n+1}^{N}\sum_{j=m+1}^{M} g(\hat{x}_{ij})m(x_{ij}) + m_1(\hat{x}_{ij}) + m_2(\hat{x}_{ij})\Bigg\}\Bigg\}.$$

From equation (33),

$$m_2(\hat{x}_{ij}) = \frac{1}{mnb}\sum_{i=n+1}^{N}\sum_{j=m+1}^{M} K\big(\frac{X_{ij}-x_{ij}}{\lambda_{ij}b}\big)e_{ij} \tag{49}$$

Hence

$$Var\sum_{i=n+1}^{N}\sum_{j=m+1}^{M}\big[m_2(\hat{x}_{ij})\big] = \frac{1}{(Mn)^2}\Bigg(\frac{Mn-mn}{mn(\lambda_{ij}b)}\Bigg)^2\sum_{i=1}^{n}\sum_{j=1}^{m}Var(D_x), \tag{50}$$

where $D_x = K\big(\frac{X_{ij}-x_{ij}}{\lambda_{ij}b}\big)e_{ij}$. Expressing equation (50) in terms of expectation the following equation is obtained

$$Var\sum_{i=n+1}^{N}\sum_{j=m+1}^{M}\big[m_2(\hat{x}_{ij})\big] = \frac{1}{(Mn)^2}\Bigg[\frac{(Mn-mn)^2}{mn(\lambda_{ij}b)^2}\Bigg]\Big\{E[D_x]^2 - [E(D_x)]^2\Big\}. \tag{51}$$

Using the fact that the conditional expectation $E(e_{ij}/X_{ij}) = 0$, the second term in equation (51) reduces to zero. Therefore,

$$Var\sum_{i=n+1}^{N}\sum_{j=m+1}^{M}\big[m_2(\hat{x}_{ij})\big] = \frac{1}{(Mn)^2}\Bigg[\frac{(Mn-mn)^2}{mn(\lambda_{ij}b)^2}\Bigg]\sigma_{ij}^2, \tag{52}$$

where $E\big(e_{ij}/X_{ij}\big)^2 = \sigma_{ij}^2$.
Let $X = X_{ij}$, and $x = x_{ij}$ and make the following substitutions

$$\left.\begin{array}{c} w = \frac{X-x}{\lambda_{ij}b} \\ X - x = (\lambda_{ij}b)w \\ dX = (\lambda_{ij}b)dw. \end{array}\right\} \tag{53}$$

so that

$$Var\sum_{i=n+1}^{N}\sum_{j=m+1}^{M}\big[m_2(\hat{x}_{ij})\big] = \frac{(Mn-mn)^2}{mn(\lambda_{ij}b)^2(Mn)^2}\int K\Bigg(\frac{X-x}{\lambda_{ij}b}\Bigg)^2\sigma_x^2\, g(X)dX \tag{54}$$

Using the change of variables technique and simplifying, equation (54) reduces to

$$Var\sum_{i=n+1}^{N}\sum_{j=m+1}^{M}\big[m_2(\hat{x}_{ij})\big] = \frac{(Mn-mn)^2}{mn(\lambda_{ij}b)(Mn)^2}\int K(w)^2\sigma_x^2\, g(x)dw + o\big(\frac{1}{mn(\lambda_{ij}b)}\big). \tag{55}$$

Following the same procedure for getting the variance of $m_2(\hat{x}_{ij})$, $Var\sum_{i=n+1}^{N}\sum_{j=m+1}^{M}\left[m_1(\hat{x}_{ij})\right]$ can similarly be obtained as follows:
(56)

$$Var\sum_{i=n+1}^{N}\sum_{j=m+1}^{M}\left[m_1(\hat{x}_{ij})\right]=\frac{1}{(Mn)^2}Var\sum_{i=n+1}^{N}\sum_{j=m+1}^{M}\left[\frac{1}{mn(\lambda_{ij}b)}\sum_{i=1}^{n}\sum_{j=1}^{m}K\left(\frac{X_{ij}-x_{ij}}{\lambda_{ij}b}\right)\right]$$
$$\times\left[m(X_{ij})-m(x_{ij})\right]$$

Equation (56) can be re-written as

$$Var\sum_{i=n+1}^{N}\sum_{j=m+1}^{M}\left[m_1(\hat{x}_{ij})\right]=\frac{(Mn-mn)^2}{mnb^2(Mn)^2}VarK\left(\frac{X_{ij}-x_{ij}}{\lambda_{ij}b}\right)^2\Big[m(X_{ij})-m(x_{ij})\Big]^2g(X)dX \tag{57}$$

where $X=(\lambda_{ij}b)w+x$ so that $dX=(\lambda_{ij}b)dw$. Changing variables and applying Taylor's series expansion about the point x_{ij} leads to

$$Var\sum_{i=n+1}^{N}\sum_{j=m+1}^{M}\Big[m_1(\hat{x}_{ij})\Big]=\frac{(Mn-mn)^2}{mn(\lambda_{ij}b)^2(Mn)^2}\int K(w^2)\Big[m(x+(\lambda_{ij}b)w)-m(x)\Big]^2\times g(x+(\lambda_{ij}b)w)dw \tag{58}$$

which gives

$$Var\sum_{i=n+1}^{N}\sum_{j=m+1}^{M}\left[m_1(\hat{x}_{ij})\right]=\frac{(Mn-mn)^2}{mn(\lambda_{ij}b)^2(Mn)^2}\int K(w^2)\Big[m(x)+m\prime(x)(\lambda_{ij}b)w+\cdots-m(x)\Big]^2\Big(g(x)+g\prime(x)(\lambda_{ij}b)w\Big)dw. \tag{59}$$

Following the procedure by (*cf.* [3]) and simplifying, equation (59) reduces to

$$Var\sum_{i=n+1}^{N}\sum_{j=m+1}^{M}\left[m_1(\hat{x}_{ij})\right]=o\left[\frac{(Mn-mn)^2b^2}{mn(\lambda_{ij}b)}\right]. \tag{60}$$

For large samples, as $n\to N, m\to M$ and $b\to 0$, then $mn(\lambda_{ij}b)\to\infty$. Hence the variance in equation (59) asymptotically tends to zero, i.e, $Var\sum_{i=1}^{N}\sum_{j=1}^{M}\left[m_1(\hat{x}_{ij})\right]\to 0$ so that the variance of the estimator of the population mean reduces to

$$Var\left(\hat{\bar{\bar{Y}}}_{INW}/x_{ij}\right)=\frac{(Mn-mn)^2}{mn(\lambda_{ij}b)(Mn)^2}\sum_{i=n+1}^{N}\sum_{j=m+1}^{M}Var\left[m(x_{ij})+\frac{m_1(\hat{x}_{ij})+m_2(\hat{x}_{ij})}{g(\hat{x}_{ij})}\right]. \tag{61}$$

Simplifying equation (61) leads to

$$Var\left(\hat{\bar{\bar{Y}}}_{INW}/x_{ij}\right)=\frac{(Mn-mn)^2}{mn(\lambda_{ij}b)(Mn)^2\big[g(\hat{x}_{ij})\big]^2}Var\left\{\sum_{i=n+1}^{N}\sum_{j=m+1}^{M}\left[m_2(\hat{x}_{ij})\right]\right\}. \tag{62}$$

Substituting equation (55) into (62) yields the following:

$$Var\left(\hat{\bar{\bar{Y}}}_{INW}/x_{ij}\right)=\frac{1}{(Mn)^2}\left\{\frac{(Mn-mn)^2\int K(w)^2\sigma^2_{x_{ij}}dw}{mn(\lambda_{ij}b)g(\hat{x}_{ij})}+o\left[\frac{(Mn-mn)^2}{mn(\lambda_{ij}b)}+\frac{1}{mn(\lambda_{ij}b)}\right]\right\}. \tag{63}$$

2.4. **Mean Squared Error of the Proposed Estimator.** The conditional MSE of the estimator of finite population mean combines the conditional squared bias and the conditional variance of the estimator, that is,

$$MSE(\hat{\bar{\bar{Y}}}_{INW}/x_{ij})=Var(\hat{\bar{\bar{Y}}}_{INW}/x_{ij})+Bias^2(\hat{\bar{\bar{Y}}}_{INW}/x_{ij}) \tag{64}$$

Which on simplification leads to

$$
\begin{aligned}
MSE(\hat{\bar{\bar{Y}}}/X_{ij} = x_{ij}) = \frac{1}{(Mn)^2}\Bigg\{ & \frac{(Mn-mn)^2 H(w)\sigma^2_{x_{ij}}}{mn(\lambda_{ij}b)g(\hat{x}_{ij})} \\
& + \Bigg[\frac{(Mn-mn)^2}{4(mn)^2(Mn^2)}(\lambda_{ij}b)^2 d_k^2\Big[m\prime\prime(x_{ij})g(x_{ij}) + \frac{2g\prime(x_{ij})m\prime(x_{ij})}{g(\hat{x}_{ij})}\Big]^2 \\
& + o\Bigg(\frac{1}{Mn}\Big\{\frac{(Mn-mn)^2}{mn(\lambda_{ij}b)} + \frac{1}{mn(\lambda_{ij}b)}\Big\}\Bigg)\Bigg]\Bigg\}
\end{aligned}
\tag{65}
$$

where $H(w) = \int K(w)^2 dw$, $d_k = \int w^2 K(w)dw$.
From equation (65), it is noted that if the sample size is large, that is as $n \to N$ and $m \to M$, the MSE of $\hat{\bar{\bar{Y}}}_{INW}$ due to the kernel tends to zero for a sufficiently small bandwidth. The estimator $\hat{\bar{\bar{Y}}}$ is therefore asymptotically consistent since its MSE converges to zero in probability.

3. Simulation Study

A simulation experiment was conducted using R code in order to compare the performance of the proposed estimator in two-stage cluster sampling with the transformed estimator due to (*cf.* [5]) and the non-parametric regression estimator due to (*cf.* [6]). An asymptotic framework is used where both the population number of clusters and the sample number of clusters are large. The number of clusters within each cluster, N_i, is held constant so that no cluster dominates the population.
Both linear and non-linear mean functions of auxiliary random variables due to (*cf.* [6]) were considered in generating data, where $x \in (0, 1)$. The equations of the mean functions used in simulating the data are given in table 1 below.

Table 1. Equations of Mean Functions Simulated

Mean function	Equation
Linear	$m_1(\hat{x}) = 1 + 2(x - 0.5)$
Quadratic	$m_2(\hat{x}) = 1 + 2(x - 0.5)^2$
Sine	$m_3(\hat{x}) = 2 + sin(2\pi x)$
Exponential	$m_4(\hat{x}) = exp(-8x)$
Bump	$m_5(\hat{x}) = 1 + 2(x - 0.5) + exp\{-200(x - 0.5)^2\}$
Jump	$m_6(\hat{x}) = 1 + 2(x - 0.5)I_{x\leq 0.65} + 0.65I_{x\geq 0.65}$

The population auxiliary values, x_{ij}, of size $M = 2000$ are generated as identical and independently distributed uniform $(0, 1)$ random variables. The survey values are only known for the respondents in the selected sample. Using the auxiliary values, the non-response values are generated, that is, for every generated value $x_{ij}, i = 1, 2, \cdots, M; j = 1, 2, \cdots, N_i$, the mean survey non-response values are generated as

$$\hat{y}_{ij} = \frac{1}{M}\Big\{\frac{m(\hat{x}_{ij})}{N_i} + \frac{\hat{e}_{ij}}{N_i}\Big\} \tag{66}$$

where $\hat{e}_{ij}$ are identically and independently distributed normal random variables with mean zero and variance one. Besides, a Gaussian kernel with mean zero and variance one was used. A Gaussian kernel was used since it has smooth and continuous derivatives at every data point. Besides, an optimal bandwidth generated using cross-validation technique due to (*cf.* [14]) was used. It has been noted by (*cf.* [14]) that this bandwidth would lead to more informative estimates compared to other choices. The local bandwidth, λ_{ij}, given in equation (**??**) were generated using the algorithm due to (*cf.* [12]).

At stage one, a sample of clusters is generated first by simple random sampling using a sample of size $m = 200$. At stage two, sub-samples of elements within every selected cluster are generated by simple random sampling with replacement using a random sample of size n_i. The non-response mean survey values were then generated using equation (66). The estimates of finite population mean were then computed using the estimator in equation (15). The values of bias and mean squared error values were also computed. The 95%

confidence intervals were then constructed for the estimators of the finite population means for comparative purposes.

4. Simulation Results

The values of the bias, mean squared error and confidence interval lengths are given in the following tables. Note that $\hat{\hat{\bar{Y}}}_{INW}$ is the estimator of finite population mean proposed in this study, $\hat{\hat{\bar{Y}}}_{TDM}$ is the transformation of data method estimator of finite population mean due to (*cf.* [5]) whereas $\hat{\hat{\bar{Y}}}_{REG}$ is the non-parametric regression estimator due to (*cf.* [6]). Both $\hat{\hat{\bar{Y}}}_{TDM}$ and $\hat{\hat{\bar{Y}}}_{REG}$ were used for comparative purposes with the proposed estimator.

Table 2. Summary Results of Bias

Estimators	$\hat{\hat{\bar{Y}}}_{INW}$	$\hat{\hat{\bar{Y}}}_{TDM}$	$\hat{\hat{\bar{Y}}}_{REG}$
Linear	-0.00213	-0.1667	-0.3312
Quadratic	-0.0132	0.04171	-0.0966
Sine	-0.0521	-0.6416	-1.2311
Exponential	-0.0041	0.3592	0.7225
Bump	-0.0032	-0.2358	-0.4685
Jump	-0.0188	-0.2466	-0.4743

The biases of the estimators considered are presented in table 2 above. Negative values of the bias imply underestimation while positive values of the bias indicate overestimation of the finite population mean by the different estimators. The proposed estimator has relatively smaller values of the bias followed by transformation of data method estimator due to (*cf.* [5]). The non-parametric-based estimator due to (*cf.* [6]) has larger values compared to the other two estimators. It is also observed that the three estimators have relatively closer values of the bias in the quadratic mean function though the transformation of data method has positive bias at this mean function. Generally, among the three estimators of finite population mean, the proposed estimator using improved Nadaraya-Watson kernel regression technique performs better than the other two estimators in terms of bias.

Table 3. Summary Results of MSE Values

Estimators	$\hat{\hat{\bar{Y}}}_{INW}$	$\hat{\hat{\bar{Y}}}_{TDM}$	$\hat{\hat{\bar{Y}}}_{REG}$
Linear	0.0334	0.1097	0.1321
Quadratic	0.0093	0.1455	0.5835
Sine	0.4215	1.5157	1.555
Exponential	0.3430	0.5220	1.3780
Bump	0.0634	0.2195	0.2508
Jump	0.2250	0.2951	1.1611

Mean squared error combines both the variance and the squared bias terms of an estimator. The mean squared error values presented in tables 3 were simulated using the different mean functions indicated. The quadratic mean function gives the smallest value of the mean squared error of the proposed estimator followed by the linear function. The estimator due to (*cf.* [6]) has the largest value of the mean squared error in the jump function. Generally, it is noted from table 3 that the mean squared error values for the proposed estimator are relatively smaller than the rest of the estimators considered. The transformation of data method estimator due to (*cf.* [5]) follows closely in the second place with smaller mean squared error values compared to non-parametric regression-based estimator due to (*cf.* [6]). From this comparison of the mean squared error values, it can be concluded that the proposed estimator is more efficient than the other two estimators considered. It has got smaller MSE values in all the mean functions and thus outperforms the others in terms of efficiency.

TABLE 4. Summary Results of 95% CI Lengths

Estimators	$\hat{\hat{\hat{Y}}}_{INW}$	$\hat{\hat{\hat{Y}}}_{TDM}$	$\hat{\hat{\hat{Y}}}_{REG}$
Linear	0.7164	1.2984	1.4249
Quadratic	0.8270	1.4951	2.994
Sine	1.5269	2.5451	4.888
Exponential	2.2958	2.8322	4.6016
Bump	0.9872	1.8365	1.9630
Jump	1.8594	2.1297	4.2239

The 95% upper and lower confidence intervals were constructed for the estimators of finite population mean. Confidence interval lengths were then obtained. The results are given in table 4. From the values obtained, it is noted that the confidence interval lengths for the proposed estimator are much tighter than those of the estimators due to (*cf.* [6]) and (*cf.* [5]). Hence, at 95% level of confidence, the estimator proposed in this study performs better than its rival estimators.

5. Conclusion

This study has developed an estimator of finite population mean in two-stage cluster sampling assuming random non-response occurs in the survey variable in the second stage of cluster sampling. Complete auxiliary information is assumed to be available in both stage one and stage two of cluster sampling. Kernel weights developed using improved Nadaraya-Watson regression technique were used in the estimation process. The theoretical properties of the proposed estimator such as asymptotic bias, variance and mean squared error were derived. Simulation results show that the proposed estimator has smaller values of the bias, smaller mean squared error values and tighter confidence interval lengths compared to the other estimators. Therefore, the estimator of finite population mean proposed in this study dominates the estimators due to (*cf.* [6]) and (*cf.* [5]) respectively.

Acknowledgements

I thank Strathmore University through the Strathmore Institute of Mathematical Sciences (SIMS) for the financial support to attend and participate in the workshop in Turkey. Secondly, my profound gratitude goes to Kenya Education Network (KENET) for paying my air-ticket to and for Turkey for the workshop.

References

[1] Ahmed, MS and Al-Titi, Omar and Al-Rawi, Ziad and Abu-Dayyeh, Walid, (2006) Estimation of a population mean using different imputation methods, Statistics Netherlands, (1247).

[2] Andridge, Rebecca R and Little, Roderick JA, (2010) A review of hot deck imputation for survey non-response, International statistical review, Wiley Online Library, 78(1), 40-64.

[3] Bach, Eric and Shallit, Jeffrey Outlaw and Shallit, Jeffrey and Jeffrey, Shallit, (1996) Algorithmic number theory: Efficient algorithms, MIT Press, (1).

[4] Bethlehem, Jelke., (2012)Using response probabilities for assessing representativity, Statistics Netherlands, Discussion Paper.

[5] Bii, Nelson Kiprono and Onyango, Christopher Ouma and Odhiambo, John, (2019) Boundary Bias Correction Using Weighting Method in Presence of Nonresponse in Two-Stage Cluster Sampling, Journal of Probability and Statistics, Hindawi.

[6] Breidt, F. Jay and Opsomer, Jean D. et al., (2000) Local polynomial regression estimators in survey sampling, The annals of statistics, Institute of Mathematical Statistics, 28(4), 1026-1053.

[7] Demir, Serdar and Toktamiş, Öniz, (2010) On the adaptive nadaraya-watson kernel regression estimators, Hacettepe Journal of Mathematics and Statistics, Hindawi, 39(3).

[8] Dorfman, Alan H., (1992) Non-parametric regression for estimating totals in finite populations, Proceedings of the Section on Survey Research Methods, American Statistical Association Alexandria, VA 622-625.

[9] Durrant, Gabriele B. et al., (2005) Imputation methods for handling item-non-response in the social sciences: a methodological review, ESRC National Centre for Research Methods and Southampton Statistical Sciences Research Institute. NCRM Methods Review Papers NCRM/002.

[10] Fife-Schaw, Chris, (2000) Surveys and sampling issues, Research methods in psychology, American Statistical Association Alexandria, (2), 88-104.

[11] Kim, Jae Kwang and Kim, Jay J, (2007) Non-response weighting adjustment using estimated response probability, Canadian Journal of Statistics, Wiley Online Library, 35(4), 501-514.
[12] Dehnad, Khosrow, (1987) Density estimation for statistics and data analysis, Taylor & Francis.
[13] Liang, Kung-Yee and Zeger, Scott L, (1993) Silverman, BW: Density Estimation for Statistics and Data Analysis, Annual review of public health, Annual Reviews 4139 El Camino Way, PO Box 10139, Palo Alto, CA 94303-0139, USA, 14(1), 43-68.
[14] Loader, Clive R. et al., (1999) Bandwidth selection: classical or plug-in?, The Annals of Statistics, Institute of Mathematical Statistics, 27(2), 415-438.
[15] Nadaraya, Elizbar A., (1964) On estimating regression, Theory of Probability & Its Applications, SIAM, 141-142.
[16] Pike, Gary R., (2008) Using weighting adjustments to compensate for survey non-response, Research in Higher Education, Springer, 49(2), 153-171.
[17] Rancourt, E and Lee, H and Särndal, CE, (1994), Bias corrections for survey estimates from data with ratio imputed values for confounded non-response, Survey Methodology, (20), 137-147.
[18] Rao, JNK, (1996) On variance estimation with imputed survey data, Journal of the American Statistical Association, Taylor & Francis, 91(434), 499-506.
[19] Ruppert, David and Sheather, Simon J. and Wand, Matthew P., (1995) An effective bandwidth selector for local least squares regression, Journal of the American Statistical Association, Taylor & Francis Group, 90(432), 1257-1270.
[20] Kreuter, Frauke and Olson, Kristen and Wagner et al., (2010) Using proxy measures and other correlates of survey outcomes to adjust for non-response: examples from multiple surveys, Journal of the Royal Statistical Society Series A: Statistics in Society, Oxford University Press, 173(2), 389-407
[21] Singh, Sarjinder, Horn, Stephen, (2000) Compromised imputation in survey sampling, Metrika, Springer, 51(3), 267-276.
[22] Takezawa, Kunio, (2005) Introduction to nonparametric regression, John Wiley & Sons, (606).
[23] Van Buuren, Stef and Boshuizen, Hendriek C and Knook, Dick L., (1999) Multiple imputation of missing blood pressure covariates in survival analysis, Statistics in medicine, Wiley Online Library, 18(6), 681-694.
[24] Watson, Geoffrey S, (1964) Smooth regression analysis, Sankhyā: The Indian Journal of Statistics, Series A, JSTOR, 359-372.

Nelson Kiprono Bii[1*] corresponding author
Strathmore University, P.O. Box 59857-00200, Nairobi, Kenya.
Christopher Ouma Onyango[2]
Kenyatta University, P.O. Box 43844-00100, Nairobi, Kenya

Email(s): nkiprono@strathmore.edu[1*] (corresponding author), christopher@ku.ac.ke[2]

A note on generation of all Pythagorean triples

Raymond Calvin Ochieng*[1], Maurice Owino Oduor[2], Vitalis Onyango-Otieno[3]

ABSTRACT. There exist several techniques used to generate Pythagorean triples. The most effective formula for generating Pythagorean triples is the Euclid's formula. Whereas the Euclid's formula generate infinitely many Pythagorean triples, it does not generate all of them. For instance, the Euclid's formula generates the triple $(3,\ 4,\ 5)$ but does not generate $(4,\ 3,\ 5)$, in which case a transposition is needed. In addition, the triple $(9,\ 12,\ 15)$ cannot be generated directly from the Euclid's formula but rather a multiplier to the triple $(3,\ 4,\ 5)$ does so. In this note, we establish a formula which generates all Pythagorean triples, primitive and non-primitive, without using a transformation and without using a multiplier. .

2020 MATHEMATICS SUBJECT CLASSIFICATIONS: 11D09, 14G50, 11A99

KEYWORDS: Primitive Pythagorean triples, non-primitive Pythagorean triples, multipliers, transpositions, cryptography.

1. INTRODUCTION

A Pythagorean Triple (PT) is a triple of positive integers $(a,\ b,\ c)$, which satisfies the Pythagorean equation

$$a^2 + b^2 = c^2,$$

where c represents the length of the hypotenuse, a and b represent the lengths of the other two sides (called legs) of a right triangle. In other words a Pythagorean triple represents the lengths of the sides of a right triangle where all the three sides have integer lengths. We say a Pythagorean triple $(a,\ b,\ c)$ is primitive if the numbers $a,\ b$ and c are pairwise co-prime [6].

We first note the parity of $a,\ b$ and c in primitive triples, that is their values modulo 2. Since $0^2 \equiv 0,\ 1^2 \equiv 1, 2^2 \equiv 0$, and $3^2 \equiv 1 \mod 4$, the only squares $modulo 4$ are 0 and 1. Letting $A = a^2,\ B = b^2$, and $C = c^2$, we have the following solutions to $A + B \equiv C \mod 4:$

$$0 + 0 \equiv 0;\ 0 + 1 \equiv 1;\ 1 + 0 \equiv 1. \tag{1}$$

($1+1 \equiv 2$ is not a solution because 2 is not a square modulo 4.) The first of these solutions corresponds only to non-primitive triples where $a,\ b$ and c are all divisible by 2. Therefore any primitive triple corresponds to one of the other two solutions. In either case, C is odd, and exactly one of A and B is always odd. Thus c is odd, and exactly one of a and b is odd.

Many formulas for generating Pythagorean triples with particular properties have been since the time of Euclid and also by several other scholars, see [1, 2, 3, 5, 8, 9, 10, 12]. In [4], integer solutions to the Pythagorean equation are determined by finding positive integers $r,\ s$ and t such that $r^2 = 2st$. The proof of this method is given in [14] and it uses a bijection which entails transforming some grid squares into an object that uniquely determines a Pythagorean triple.

In [11], McCullough uses the height and excess enumeration to generate Pythagorean triples. In this method, for a Pythagorean triple $(a,\ b,\ c)$, the height h is just $c - b$, and the excess e is $a + b - c$. The term excess arises from the fact that e is simply the extra distance one must travel when going along the two legs instead of the hypotenuse. In [13], Roy and Sonia uses the difference between one leg and hypotenuse to formulate a method of generating Pythagorean triples.

Euclid's formula presents the most common way of generating Pythagorean triples. It states that primitive Pythagorean triples $(a,\ b,\ c)$ in which b is even, are generated by the formulae

$$a = n^2 - m^2, \quad b = 2nm, \quad c = m^2 + n^2,$$

where $n > m > 0;\ m, n \in \mathbb{Z}^+$, for any pair of co-prime positive integers of opposite parity. A triple generated using this method is primitive if $(n,\ m) = 1$ and if n and m are of opposite parity [6, 15, 16]. In this formula, a convention is made in which a is always odd and b is always even, since otherwise we can rename the variables of a given triple to obtain this. Therefore, using this formula, a Primitive Pythagorean Triple has a unique representation $(a,\ b,\ c)$, where b is even and a and c are odd. To obtain the other set of solutions $(b,\ a,\ c)$, a transformation is required. We address this problem in the subsequent sections.

2. Unit Circle

In this section, we use a parametrization of the Unit Circle to find rational points on the Unit Circle (point $(x,\ y)$ satisfying the equation $x^2 + y^2 = 1$, where both x and y are rational numbers).

Consider the lines passing through the point $(-1,\ 0)$ on the Unit Circle. Let $t = m$ be the slope of one such a line. Notice that every line through $(-1,\ 0)$ intersects the circle at exactly one point. Next, we point out that since the distance from $(-1,\ 0)$ to the origin is 1, the line with slope t will have a $y-$intercept of $(0,\ t)$. Thus, except the vertical line $x = -1$, the other lines have equations of the form

$$y = tx + t. \tag{2}$$

Moreover, each of these lines cuts the circle in exactly two points, one of which is $(-1,\ 0)$. Let us find the other point.

Substituting (2) into the equation of the Unit Circle gives

$$x^2 + t^2(x+1)^2 = 1,$$

or

$$(1+t^2)x^2 + 2t^2x + (t^2 - 1) = 0. \tag{3}$$

Solving (3) for x, we find

$$\begin{aligned} x &= \frac{-2t^2 \pm \sqrt{(4t^4 - 4(1+t^2)(t^2-1))}}{2(1+t^2)} \\ &= -1 \ or \ \frac{1-t^2}{1+t^2} \end{aligned}$$

and $y = t(x+1) = 0$ or $t(\frac{1-t^2}{1+t^2} + 1) = \frac{2t}{1+t^2}$.

Thus, we have a parametrization of the Unit Circle, excluding the point $(-1,\ 0)$, with parameter t :

$$x(t) = \frac{1-t^2}{1+t^2}; \qquad y(t) = \frac{2t}{1+t^2}. \tag{4}$$

It can easily be checked that this is indeed the Unit Circle by substituting into the Cartesian equation $x^2 + y^2 = 1$.

Substituting any rational number for the parameter t in (4) will give a rational number. For example, if $t = 1$, then $x = 0$ and $y = 1$. If $t = \frac{1}{2}$, then $x = \frac{3}{5}$, $y = \frac{4}{5}$ and if $t = \frac{1}{5}$ then

$x = \frac{24}{26} = \frac{12}{13}$, $y = \frac{10}{26} = \frac{5}{13}$, and so on. Each rational point yields a Pythagorean Triple. For example,

$$(\frac{3}{5})^2 + (\frac{4}{5})^2 = 1$$

implies that $(3,\ 4,\ 5)$ is a Pythagorean Triple;

$$(\frac{24}{26})^2 + (\frac{10}{26})^2 = 1$$

implies that $(24,\ 10,\ 26)$ is a Pythagorean Triple; and the simplified form of this equation

$$(\frac{12}{13})^2 + (\frac{5}{13})^2 = 1$$

yields the Pythagorean Triple $(12,\ 5,\ 13)$.

Pythagorean Triples

Let $(x, y) = (\frac{1-t^2}{1+t^2}, \frac{2t}{1+t^2})$ be a parametrization of the Unit Circle $x^2 + y^2 = 1$, where t is the slope of any line through the point $(-1,\ 0)$ and cutting the circle at $(x,\ y)$. For any t with $0 < t < 1$, the parametric equations $x(t)$ and $y(t)$ yield Pythagorean Triples $(a,\ b,\ c)$. In the table below, we enumerate some Triples for $t = \frac{r}{s} < 1$, where $0 < r < s < 10$, and $(r,\ s) = 1$.

s	r	t	x	y	$PT = (a,\ b,\ c)$	$PPT = (a,\ b,\ c)$
2	1	$\frac{1}{2}$	$\frac{3}{5}$	$\frac{4}{5}$	$(3,\ 4,\ 5)$	$(3,\ 4,\ 5)$
3	1	$\frac{1}{3}$	$\frac{8}{10}$	$\frac{6}{10}$	$(8,\ 6,\ 10)$	$(4,\ 3,\ 5)$
3	2	$\frac{2}{3}$	$\frac{5}{13}$	$\frac{12}{13}$	$(5,\ 12,\ 13)$	$(5,\ 12,\ 13)$
4	1	$\frac{1}{4}$	$\frac{15}{17}$	$\frac{8}{17}$	$(15,\ 8,\ 17)$	$(15,\ 8,\ 17)$
4	3	$\frac{3}{4}$	$\frac{7}{25}$	$\frac{24}{25}$	$(7,\ 24,\ 25)$	$(7,\ 24,\ 25)$
5	1	$\frac{1}{5}$	$\frac{24}{26}$	$\frac{10}{26}$	$(24,\ 10,\ 26)$	$(12,\ 5,\ 13)$
5	2	$\frac{2}{5}$	$\frac{21}{29}$	$\frac{20}{29}$	$(21,\ 20,\ 29)$	$(21,\ 20,\ 29)$
5	3	$\frac{3}{5}$	$\frac{16}{34}$	$\frac{30}{34}$	$(16,\ 30,\ 34)$	$(8,\ 15,\ 17)$
6	1	$\frac{1}{6}$	$\frac{35}{37}$	$\frac{12}{37}$	$(35,\ 12,\ 37)$	$(35,\ 12,\ 37)$
6	5	$\frac{5}{6}$	$\frac{11}{61}$	$\frac{60}{61}$	$(11,\ 60,\ 61)$	$(11,\ 60,\ 61)$
7	1	$\frac{1}{7}$	$\frac{48}{50}$	$\frac{14}{50}$	$(48,\ 14,\ 50)$	$(24,\ 7,\ 25)$
7	2	$\frac{2}{7}$	$\frac{45}{53}$	$\frac{28}{53}$	$(45,\ 28,\ 53)$	$(45,\ 28,\ 53)$
7	3	$\frac{3}{7}$	$\frac{40}{58}$	$\frac{42}{58}$	$(40,\ 42,\ 58)$	$(20,\ 21,\ 29)$
7	4	$\frac{4}{7}$	$\frac{33}{65}$	$\frac{56}{65}$	$(33,\ 56,\ 65)$	$(33,\ 56,\ 65)$
8	1	$\frac{1}{8}$	$\frac{63}{65}$	$\frac{16}{65}$	$(63,\ 16,\ 65)$	$(63,\ 16,\ 65)$
8	3	$\frac{3}{8}$	$\frac{55}{73}$	$\frac{48}{73}$	$(55,\ 48,\ 73)$	$(55,\ 48,\ 73)$
8	5	$\frac{5}{8}$	$\frac{39}{89}$	$\frac{80}{89}$	$(39,\ 80,\ 89)$	$(39,\ 80,\ 89)$
8	7	$\frac{7}{8}$	$\frac{15}{113}$	$\frac{112}{113}$	$(15,\ 112,\ 113)$	$(15,\ 112,\ 113)$
9	1	$\frac{1}{9}$	$\frac{80}{82}$	$\frac{18}{82}$	$(80,\ 18,\ 82)$	$(40,\ 9,\ 41)$
9	2	$\frac{2}{9}$	$\frac{77}{85}$	$\frac{36}{85}$	$(77,\ 36,\ 85)$	$(77,\ 36,\ 85)$
9	4	$\frac{4}{9}$	$\frac{65}{97}$	$\frac{72}{97}$	$(65,\ 72,\ 97)$	$(65,\ 72,\ 97)$
9	5	$\frac{5}{9}$	$\frac{56}{106}$	$\frac{90}{106}$	$(56,\ 90,\ 106)$	$(28,\ 45,\ 53)$
9	7	$\frac{7}{9}$	$\frac{32}{130}$	$\frac{126}{130}$	$(32,\ 126,\ 130)$	$(16,\ 63,\ 65)$
9	8	$\frac{8}{9}$	$\frac{17}{145}$	$\frac{144}{145}$	$(17,\ 144,\ 145)$	$(17,\ 144,\ 145)$
10	1	$\frac{1}{10}$	$\frac{99}{101}$	$\frac{20}{101}$	$(99,\ 20,\ 101)$	$(99,\ 20,\ 101)$
10	3	$\frac{3}{10}$	$\frac{91}{109}$	$\frac{60}{109}$	$(91,\ 60,\ 109)$	$(91,\ 60,\ 109)$
10	7	$\frac{7}{10}$	$\frac{51}{149}$	$\frac{140}{149}$	$(51,\ 140,\ 149)$	$(51,\ 140,\ 149)$
10	9	$\frac{9}{10}$	$\frac{19}{181}$	$\frac{180}{181}$	$(19,\ 180,\ 181)$	$(19,\ 180,\ 181)$

Table 1. Triples for values of t with $0 < r < s < 10$.

Note 2.1. *We note that this method is incapable of generating all the Pythagorean triples. For instance, the triple* $(4,\ 3,\ 5)$ *cannot be generated directly. A transposition on the triple* $(3,\ 4,\ 5)$ *is applied to obtain* $(4,\ 3,\ 5)$*. In addition the triple* $(9,\ 12,\ 15)$ *requires a multiplier on* $(3,\ 4,\ 5)$*. We establish a formula that generates all Pythagorean triples without need for either transpositions or multipliers.*

Theorem 2.2. *Let* $(a,\ b,\ c)$ *be a primitive Pythagorean triple with* b *even. Then there exist relatively prime positive integers* $m < n$ *having distinct parities (i.e. one even and one odd) such that*

$$a = n^2 - m^2, \quad b = 2nm, \quad c = n^2 + m^2. \tag{5}$$

In Table 2, we generate a few examples of Pythagorean triples using equation (5). We can obtain an equivalent formula to Theorem 2.2 from the equation $a^2 + b^2 = c^2$ by writing

$$a^2 = c^2 - b^2 = (c - b)(c + b)$$

in place of $b^2 = c^2 - a^2$, for a given primitive Pythagorean triple $(a,\ b,\ c)$.

n	m	$a = n^2 - m^2$	$b = 2nm$	$c = n^2 + m^2$	$DL(c,\ a) = 2m^2$	$DL(c,\ b) = u^2$
2	1	3	4	5	2	1
3	1	8	6	10	2	4
3	2	5	12	13	8	1
4	1	15	8	17	2	9
4	2	12	16	20	8	4
4	3	7	24	25	18	1
5	1	24	10	26	2	16
5	2	21	20	29	8	9
5	3	16	30	34	18	4
5	4	9	40	41	32	1

Table 2. Examples of Pythagorean triples generated from equation (5).

Theorem 2.3. *Every primitive Pythagorean triple* $(a,\ b,\ c)$ *with* a *odd and* b *even can be obtained by using the formulas*

$$a = st, \quad b = \frac{s^2 - t^2}{2}, \quad c = \frac{s^2 + t^2}{2},$$

where $s > t \geqslant 1$ *are chosen to be any odd integers with no common factors.*

Notice that Theorems 2.2 and 2.3 are equivalent under the substitution $t = n - m$, $s = n + m$.

Remark 1. *Let*

$$(a, b, c) = (n^2 - m^2,\ 2nm,\ n^2 + m^2),$$

be a primitive Pythagorean triple, where $n > m$, $(n,\ m) = 1$ *and* n *and* m *are of opposite parity.*

The two methods in Theorems 2.2 and 2.3, like many others will require a transposition. For instance in both b *is always even whenever* $(a,\ b,\ c)$ *is primitive.*

In the following section, we consider a formula which is able to generate all Pythagorean triples without need for either multipliers of transpositions.

3. Generating All Pythagorean Triples

It can be seen that the difference formula does not generate all Pythagorean triples. For instance the triple $(9,\ 12,\ 15)$ cannot be generated directly by the formula but rather applying a multiplier of 3 to the triple $(3,\ 4,\ 5)$ does so. Moreover, for every primitive triple $(a,\ b,\ c)$ generated by the formula, the corresponding primitive triple $(b,\ a,\ c)$ cannot be directly generated, a transposition is required.

If $(a,\ b,\ c)$ is a Pythagorean triple, the triple $(b,\ a,\ c)$ is called its co-triple. In this section, we illustrate an expression which generates all Pythagorean triples uniquely, as discussed below.

Proposition 3.1. *Let $(a,\ b,\ c)$ be a Pythagorean triple. Then every Pythagorean triple can be generated by the formula:*

$$(a,\ b,\ c) = \left(\sqrt{PQ},\ \frac{P-Q}{2},\ \frac{P+Q}{2}\right), \tag{6}$$

for some positive integers $P,\ Q$.

Proof. A Pythagorean triple $(a,\ b,\ c)$ satisfies the equation $a^2 + b^2 = c^2$. From this we see that $a^2 = c^2 - b^2 = (c+b)(c-b)$. Let $P = c+b$ and $Q = c-b$. Clearly $P > Q$. Solving for c and b, one obtains:

$$P + Q = 2c \quad \Leftrightarrow \quad c = \frac{P+Q}{2}$$

and

$$P - Q = 2b \quad \Leftrightarrow \quad b = \frac{P-Q}{2}.$$

Since P and Q are integers, P and Q are either both even or both odd. So for each a, a^2 is a product of two factors P and Q, with $P > Q > 0$.
Next, suppose $c = b + d$ where d is the difference between c and b, then

$$a^2 + b^2 = (b+d)^2,$$

which simplifies to

$$a^2 = 2bd + d^2$$

or

$$a^2 = d(2b+d). \tag{7}$$

Substitute $\frac{P-Q}{2}$ for b in (7) to obtain:

$$\begin{aligned} a^2 &= d[2(P-Q)+d] = PQ \\ &\Leftrightarrow dP - dQ + d^2 = PQ \\ &\Leftrightarrow dP + d^2 = (P+d)Q \\ &\Leftrightarrow d(P+d) = (P+d)Q. \end{aligned}$$

Thus $d = Q$. So $a = \sqrt{PQ}$, where $P > Q$ and positive integers $P,\ Q$ are either both odd or both even. Then

$$(a,\ b,\ c) = \left(\sqrt{PQ},\ \frac{P-Q}{2},\ \frac{P-Q}{2}\right).$$

Next, we show that this is a Pythagorean triple, that is,

$$\begin{aligned} \left(\sqrt{PQ}\right)^2 + \left(\frac{P-Q}{2}\right)^2 &= PQ + \frac{P^2 - 2PQ + Q^2}{4} \\ &= \frac{P^2 + 2PQ + Q^2}{4} \\ &= \left(\frac{P+Q}{2}\right)^2. \end{aligned}$$

□

For every primitive Pythagorean triple $(a,\ b,\ c)$ one can easily obtain the corresponding co-triple $(b,\ a,\ c)$ from (6) as illustrated in the corollary below:

Corollary 3.2. *If $P = u^2$ and $Q = v^2$ where $u,\ v$ are relatively prime odd positive integers, such that $u > v$, then $(a,\ b,\ c)$ is a primitive Pythagorean triple with a odd. The corresponding co-triple is generated when $P = 2\left(\frac{u+v}{2}\right)^2$ and $Q = 2\left(\frac{u-v}{2}\right)^2$.*

Proof. Let $P = u^2$ and $Q = v^2$, then by Proposition 3.1, we obtain

$$(a,\ b,\ c) = \left(uv,\ \frac{u^2 - v^2}{2},\ \frac{u^2 + v^2}{2}\right). \tag{8}$$

Clearly, if $u > v > 0$ are odd integers such that $(u,\ v) = 1$, then (8) is a primitive Pythagorean triple with a odd.

If $P = 2\left(\frac{u+v}{2}\right)^2$ and $Q = 2\left(\frac{u-v}{2}\right)^2$, then

$$\sqrt{PQ} = \sqrt{2\left(\frac{u+v}{2}\right)^2 2\left(\frac{u-v}{2}\right)^2} = \frac{u^2 - v^2}{2} = b,$$

$$\frac{P-Q}{2} = \sqrt{\frac{2\left(\frac{u+v}{2}\right)^2 - 2\left(\frac{u-v}{2}\right)^2}{2}} = \frac{u^2 + 2uv + v^2 - u^2 + 2uv - v^2}{4} = \frac{4uv}{4} = uv = a,$$

and

$$\frac{P+Q}{2} = \sqrt{\frac{2\left(\frac{u+v}{2}\right)^2 - 2\left(\frac{u-v}{2}\right)^2}{2}} = \frac{u^2 + 2uv + v^2 + u^2 - 2uv + v^2}{4} = \frac{u^2 + v^2}{2} = c.$$

□

Non-primitive Pythagorean triples can be generated as described below:

Corollary 3.3. *If P and Q are not relatively prime, then $(a,\ b,\ c)$ corresponding $(b,\ a,\ c)$ are non-primitive Pythagorean triples.*

Proof. Let $P = ku^2$ and $Q = kv^2$, then by (6),

$$\left(\sqrt{ku^2 \cdot kv^2},\ \frac{ku^2 - kv^2}{2},\ \frac{ku^2 + kv^2}{2}\right) = k \cdot \left(uv,\ \frac{u^2 - v^2}{2},\ \frac{u^2 + v^2}{2}\right) = k(a,\ b,\ c).$$

Similarly, if $P = 2k\left[\frac{u+v}{2}\right]^2$ and $Q = 2k\left[\frac{u-v}{2}\right]^2$, then

$$\begin{aligned}
&\left(\sqrt{2k\left[\frac{u+v}{2}\right]^2 \cdot 2k\left[\frac{u-v}{2}\right]^2},\ \frac{2k\left[\frac{u+v}{2}\right]^2 - 2k\left[\frac{u-v}{2}\right]^2}{2},\ \frac{2k\left[\frac{u+v}{2}\right]^2 + 2k\left[\frac{u-v}{2}\right]^2}{2}\right) \\
= &\left(2k\left[\frac{u+v}{2}\right] \cdot \left[\frac{u-v}{2}\right],\ 2k \cdot \frac{4uv}{8},\ 2k \cdot \frac{2u^2 + 2v^2}{8}\right) \\
= &\ k \cdot \left(\frac{u^2 - v^2}{2},\ uv,\ \frac{u^2 + v^2}{2}\right) \\
= &\ k(b,\ a,\ c).
\end{aligned}$$

□

Tables 3 and 4 below shows examples of primitive and non-primitive triples and their co-triples generated by this method:

u	v	P	Q	a	b	c	k	m	P	Q	b	a	c
3	1	9	1	3	4	5	2	1	8	2	4	3	5
5	1	25	1	5	12	13	3	2	18	8	12	5	13
7	1	49	1	7	24	25	4	3	32	18	24	7	25
9	1	81	1	9	40	41	5	3	50	32	40	9	41
5	3	25	9	15	8	17	4	1	32	2	8	15	17
7	3	49	9	21	20	29	5	2	50	8	20	21	29
11	3	121	9	33	56	65	7	4	98	72	56	33	65
13	3	169	9	39	80	89	8	5	128	50	80	39	89

Table 3. Primitive Pythagorean Triples and their co-triples

P	Q	a	b	c	P	Q	b	a	c
18	2	6	8	10	16	4	8	6	10
27	3	9	12	15	24	6	12	9	15
36	4	12	16	20	32	8	16	12	20
45	5	15	20	25	40	10	20	15	25
50	18	30	16	34	64	4	16	30	34
98	18	42	40	58	100	16	40	42	58
242	18	66	112	130	196	144	112	66	130
338	18	78	160	178	256	100	160	78	178

Table 4. Non-primitive Pythagorean Triples and their co-triples

4. Applications in Cryptography

Theorem 4.1. *Primitive Pythagorean triples come in* 6 *classes based on the divisibility of* $a,\ b,\ c$ *by* 3, 4, *and* 5.

Proof. 1 Class A: a is divisible by 3 and c is divisible by 5. e.g (3,4,5), (33,56,65)
2 Class B: a is divisible by 5, and b is divisible by 3. e.g (5,12,13)
3 Class C: a is divisible by 3 and 5. e.g (15,8,17), (45,28,53)
4 Class D: b is divisible by 3 and c by 5. e.g (7,24,25), (13,84,85)
5 Class E: a is divisible by 3 and b by 5. e.g (21,20,29), (9,40,41)
6 Class F: b is divisible by 3 and 5. e.g (11,60,61), (91,60,109)

□

An experiment was done where 4448 primitive Pythagorean triples were generated by Euclid's formula, which is limited in the Pythagorean Triples it generates, and indexed by increasing $a,\ b$, and c respectively.
Indexed by increasing a:
ABDEFDCCDFEEDBAFFAABBDEEFDCCDFEEDBBAAFFAABBDEEFDCCDDFEDBB
AAFFFAABBDEEFDCCCCDFEEDBBAAFFFAABDEEFDDCCDFEEDDBBAA...
Indexed by increasing b:
ACBBAEEDDCCCDDEEAABBCCAAFFFFACCBBAAEEDDDDCCCCDDEEAABBCCAA
FFFFAACBBBBAAEEEEDDCCCCDDDDEEAABBBBCCAAFFFFAACCBBAAEEDDDD...
Indexed by increasing c:
ABCDEBECFAABDDEBEFCACDDEBFCFAABCDDEEFCFCDDEBEFCAABCDDBFC
AABCCEBEFFAABDDEBEFFAABBDDEECCFAACDDEBCFBCDEEECFBCDDEEBB
EFAABDDBEFCA...

We obtain separate sequences related to the occurrence of As, Bs, Cs, Ds, Es, and Fs by considering the distance between occurrences of the letters. Thus in the listing by increasing c, A occurs, after its first value, at the 10th, 11th , 20th , ... positions, which corresponds to the numbers 9, 1, 9, ... These sequences are called Baudhāyana sequences.

However, equation (6) makes it easy to generate and conveniently classify all triples into these classes. This helps improve the randomness property of Baudhāyana sequences which have applications in key distribution and in information hiding in the field of Cryptography. For instance, consider the triples generated by (6) and indexed by increasing a. We obtain the following

a	b	c	Class
3	4	5	A
4	3	5	D
5	12	13	B
7	24	25	D
8	15	17	F
9	40	41	E
11	60	61	F
12	5	13	E
12	35	37	E
13	84	85	D
15	8	17	C
16	63	65	D
17	144	145	D
19	180	181	F
20	21	29	B
20	99	101	B
21	20	99	E
21	220	221	E
23	264	265	D
24	7	25	A
24	143	145	A

Table 5. Primitive Pythagorean Triples by increasing values of a

Thus in the listing by increasing a, A occurs, after its first value, at the 20th, 21st which correspond to the numbers 1, 18, 1, ...
Similarly in the listing by increasing c, A occurs, after its first value, at the 8th, 18th which correspond to the numbers 1, 7, 17, ...

Conclusion

Proposition 3.1 presents an effective way of generating all Pythagorean triples without need for multipliers or transpositions. It is easy to determine the values of the generating pair of integers P and Q such that (6) is a triple. If $u > v$ are relatively prime odd positive integers such that $u^2 = P$ and $v^2 = Q$ then P, Q are odd and $(P, Q) = 1$. In this case a is odd and (a, b, c) is primitive. Similarly, if $k > m$ are any relatively prime positive integers such that $P = 2k^2$ and $Q = 2m^2$, then $(P, Q) = 2$. Clearly a is even and (a, b, c) is primitive. To obtain non-primitive triples, one may use the expression $n(P, Q)$ for $n \in \mathbb{Z}^+$.

This formula makes it easy to generate and conveniently classify triples into classes according to divisibility by 3 or 5 or both. This helps improve the randomness property and its application to Cryptography.

Acknowledgement

This work was done while the first author was conducting his PhD research at Strathmore University, and would like to thank the University for financial support.

References

[1] Beauregard, R. A. and Suryanarayan, E. R. (2000), Parametric representation of primitive Pythagorean Triples, in Nelsen, R. B., Proofs Without Words: Exercises in Visual Thinking II, Mathematical Association of America, p. 120, ISBN 978-0-88385-721-2, OCLC 807785075.

[2] Bernhart, F. R. and Price, H. L. (2005), Heron's Formula, Descartes circles and Pythagorean Triangles. arXiv:math/0701624.

[3] Carmichael, R. D. (1959), The Theory of Numbers and Diophantine Analysis, Dover Publications.

[4] Dickson, L. E. (1920), History of the Theory of Numbers, Vol.II. Diophantine Analysis, *Carnegie Institution of Washington*, Publication No. **256**, 12+803pp Read online - University of Toronto.

[5] Houston, D. (1993), 'Pythagorean Triples via double - angle formulas' in Nelsen, R. B., Proofs Without Words: Exercises in Visual Thinking, Mathematical Association of America, p. 141, ISBN 978-0-88385-700-7, OCLC 29664480.

[6] Joyce, D. E. (June 1997), "Euclid's Elements" Book X , Proposition XXIX, Clark University.

[7] Kak, S. (2010), Pythagorean Triples and Cryptographic Coding. arxiv.org/pdf/1004.3770.

[8] Koshy, T. (2002), Elementary Number Thoery with Applications, Academic Press, p. 545, ISBN 9780124211711.

[9] MacHale, Des; van den Bosch C. (2012), Generalising a result about Pythagorean triples, *Mathematical Gazette* **96**: 9196.

[10] Maor, E. (2007), The Pythagorean Theorem, Princeton University Press. New Jersey.

[11] McCullough, D. (2005), Height and Excess of Pythagorean Triangles, *Mathematics Magazine* **78**, no. 1, 26–44.

[12] Posamentier, A. S. (2010), The Pythagorean Theorem: The Story of Ots Power and Beauty, Prometheus Books, p. 156, ISBN 9781614183.

[13] Roy, T. and Sonia, F. J. (2012), A Direct Method To Generate Pythagorean Triples And Its Generalization To Pythagorean Quadruples And n-tuples, *arXiv:1201.2145 (math.NT)*, 1-11, 2012

[14] Rukavicka, J. (2013), Dickson's Method for Generating Pythagorean Triples Revisited, *European Journal of Pure and Applied Mathematics* ISSN 1307-5543, Vol. 6, No. **3** (2013) p.363-364.

[15] Sierpiski, W. (2003), Pythagorean Tiangles, Dover Publications, Inc. Mineola, New York.

[16] Spezeski, W. (2008), Rethinking Pythagorean Triples, *Applications and Applied Mathematics:* ISSN:1932-9466, Vol. 3,No. **1** (June 2008) p. 100-112

Raymond Calvin Ochieng[1*] (corresponding author)
Strathmore Institute of Mathematical Sciences, Strathmore University, Nairobi, Kenya
Maurice Owino Oduor[2]
Department of Mathematics, Actuarial and Physical Sciences, University of Kabianga, Kericho, Kenya
Vitalis Onyango-Otieno[3]
Strathmore Institute of Mathematical Sciences, Strathmore University, Nairobi, Kenya

Email(s): rochieng@strathmore.edu[1*] (corresponding author), moduor@kabianga.ac.ke[2] vonyango@strathmore.edu[3]

CMA
CONSTRUCTIVE MATHEMATICAL ANALYSIS

https://www.iwcma.selcuk.edu.tr

https://dergipark.org.tr/en/pub/cma

International Workshop Constructive Mathematical Analysis
IWCMA