

Properties of Energy-Like Statistics Under Squared Euclidean Distance and Exponential Gamma Distributions

Ahmed Jebur Ali¹, Abeer Ali Naji¹, Baneen Hussein Abd Zaid¹, Kawthar Hataf Fazaa¹ and Ali J. Ramadhan^{2*}

¹Department of Mathematics, College of Co-educational Education, University of Kufa, Najaf, Iraq

²Department of Computer Techniques Engineering, College of Technical Engineering, University of AlKafeel, Najaf, Iraq

Abstract. This study investigates distance-based statistical parameters and squared Euclidean distance for samples from exponential and gamma distributions. Statistical tests can be used as two-sample tests to determine the homogeneity between the distributions of two data sets, whether they are circular or linear. This study uses standard squared Euclidean distance linear data to calculate the properties of the proposed statistical tests, including one-dimensional and two-dimensional or higher assumptions.

1. Introduction

A two-sample permutation test is a two-independent-sample scenario used to test hypotheses about the similarities and differences between two populations or groups. It maintains a fixed sample size for each population and computes the selected test statistic for each permutation sample. The collected values form the distribution used for p-value calculations. This test is not based on an assumption about the underlying shape of the distribution, but on the null hypothesis [1].

Two-sample problems are complex and require the use of multivariate tests. Classical methods such as the Kolmogorov-Smirnov and Cramér-von Mises tests do not have a natural distribution-free extension to the multivariate case. These tests are based on distributional assumptions about the underlying population and therefore are not suitable for general two-sample or k-sample problems. To solve this problem, many methods require computational methods. Bickel developed a consistent distribution-free multivariate extension of the univariate Smirnov test, while Friedman and Rafsky proposed a distribution-free multivariate extension of the Wald-Wolfowitz test. The nearest neighbor test is another consistent, asymptotically distribution-free test for multivariate problems. Baringhaus, Franz, Székely, and Rizzo independently developed a multivariate nonparametric test for uniform distribution [2,3]. However, all the above non-parametric tests cannot be used when the data dimension is larger than the sample size, and they perform poorly when applied to high-dimensional data.

* Corresponding author: ali.j.r@alkafeel.edu.iq

The term "energy statistics" was introduced by Székely in a series of lectures in 1984-1985. Székely's basic idea and name are derived from Newton's concept of potential energy. Statistical observations can be thought of as objects in a metric space that are governed by statistical potential energy, which is zero if and only if the underlying statistical null hypothesis is true. Energy statistics (E-statistics) are a class of statistics that are a function of the distance between observations [4].

Let \mathfrak{U} and \mathfrak{X} be two independent random variables and let $\{\mathfrak{U}_1, \mathfrak{U}_2, \dots, \mathfrak{U}_\mathfrak{M}\}$ and $\{\mathfrak{X}_1, \mathfrak{X}_2, \dots, \mathfrak{X}_\mathfrak{M}\}$ be independent samples from these random variables then the energy statistic is defined by

$$\mathcal{E}(\mathfrak{U}, \mathfrak{X}) = \frac{\mathfrak{M}\mathfrak{M}}{\mathfrak{M}+\mathfrak{M}} \left(\frac{2}{\mathfrak{M}\mathfrak{M}} \mathfrak{f}_1 - \frac{1}{\mathfrak{M}^2} \mathfrak{f}_2 - \frac{1}{\mathfrak{M}^2} \mathfrak{f}_3 \right) \quad (1)$$

where

$$\mathfrak{f}_1 = \sum_{i=1}^{\mathfrak{M}} \sum_{j=1}^{\mathfrak{M}} \mathcal{D}(\mathfrak{U}_i, \mathfrak{X}_j), \quad \mathfrak{f}_2 = \sum_{i=1}^{\mathfrak{M}} \sum_{j=1}^{\mathfrak{M}} \mathcal{D}(\mathfrak{U}_i, \mathfrak{U}_j), \quad \mathfrak{f}_3 = \sum_{i=1}^{\mathfrak{M}} \sum_{j=1}^{\mathfrak{M}} \mathcal{D}(\mathfrak{X}_i, \mathfrak{X}_j),$$

and $\mathcal{D}(\mathfrak{U}_i, \mathfrak{X}_j)$ represents the distance measure between the observations \mathfrak{U}_i and \mathfrak{X}_j [2][5].

However, since the distribution of the energy statistic is unknown, this statistic can be used with a permutation test (which is very time-consuming) to test for homogeneity between the two sets of distributions. Therefore, the following question arises: Can we reduce the computational time of permutation tests by considering energy-like statistics?

In this paper, we answer this question by first considering an energy-like statistic that can be used as a test statistic to determine homogeneity between two sets of distributions. Furthermore, normal samples, squared Euclidean distance, and test characteristics in different dimensions are theoretically calculated.

2. PROPERTIES OF THE ENERGY-LIKE STATISTIC WITH ONE-DIMENSIONAL CASE ($p = 1$)

2.1-The expected value for the statistic $\mathcal{T}_{\mathfrak{R}}$ is computed in the following way: - Assume that \mathfrak{U} and \mathfrak{X} are two independent random variables such that $\mathfrak{U} \sim \text{Exp}(\lambda)$ and $\mathfrak{X} \sim \text{Exp}(\lambda)$ and that $\mathfrak{S}_1 = \{\mathfrak{U}_1, \mathfrak{U}_2, \dots, \mathfrak{U}_{\mathfrak{M}_1}\}$ and $\mathfrak{S}_2 = \{\mathfrak{X}_1, \mathfrak{X}_2, \dots, \mathfrak{X}_{\mathfrak{M}_2}\}$ are two random samples from these random variables, then

$$\begin{aligned} E(\mathcal{T}_{\mathfrak{R}}(\mathfrak{S}_1, \mathfrak{S}_2)) &= E\left(\left(\frac{\mathfrak{M}_1 + \mathfrak{M}_2}{\mathfrak{M}_1 \mathfrak{M}_2} \right)^{-\frac{1}{2}} \sum_{i=1}^{\mathfrak{M}_1} \sum_{j=1}^{\mathfrak{M}_2} \mathcal{D}(\mathfrak{U}_i, \mathfrak{X}_j) \right) \\ &= \left(\frac{\mathfrak{M}_1 + \mathfrak{M}_2}{\mathfrak{M}_1 \mathfrak{M}_2} \right)^{-\frac{1}{2}} \sum_{i=1}^{\mathfrak{M}_1} \sum_{j=1}^{\mathfrak{M}_2} E(\mathcal{D}(\mathfrak{U}_i, \mathfrak{X}_j)) \end{aligned}$$

Since $\mathcal{D}(\mathfrak{U}_i, \mathfrak{X}_j) = (\mathfrak{U}_i - \mathfrak{X}_j)^2$, $i = 1, 2, \dots, \mathfrak{M}_1$; $j = 1, 2, \dots, \mathfrak{M}_2$, then we get:

$$E(\mathcal{T}_{\mathfrak{R}}(\mathfrak{S}_1, \mathfrak{S}_2)) = \left(\frac{\mathfrak{M}_1 + \mathfrak{M}_2}{\mathfrak{M}_1 \mathfrak{M}_2} \right)^{-\frac{1}{2}} \sum_{i=1}^{\mathfrak{M}_1} \sum_{j=1}^{\mathfrak{M}_2} E((\mathfrak{U}_i - \mathfrak{X}_j)^2) \quad (2)$$

The formula $E[(\mathcal{U}_i - \mathcal{X}_j)^2]$ in Equation (2) is computed under the assumption of the independent and identically distributed (i.i.d) normal samples as follows:

$$E((\mathcal{U} - \mathcal{X})^2) = \int_0^\infty \int_0^\infty (\mathcal{U} - \mathcal{X})^2 f(\mathcal{U}, \mathcal{X}) d\mathcal{U}d\mathcal{X}$$

Given that the random variables \mathcal{U} and \mathcal{X} have similar distributions and are independent, the following results from various mathematical procedures are obtained:

$$E((\mathcal{U} - \mathcal{X})^2) = E(\mathcal{U}^2) - 2E(\mathcal{U})E(\mathcal{X}) + E(\mathcal{X}^2) = 2\lambda^2 - 2\lambda^2 + 2\lambda^2 = 2\lambda^2$$

So, Equation (2) will equal to

$$= \left(\frac{\mathfrak{Y}_1 + \mathfrak{Y}_2}{\mathfrak{Y}_1 \mathfrak{Y}_2}\right)^{-\frac{1}{2}} \sum_{i=1}^{\mathfrak{Y}_1} \sum_{j=1}^{\mathfrak{Y}_2} 2\lambda^2 = \left(\frac{\mathfrak{Y}_1 + \mathfrak{Y}_2}{\mathfrak{Y}_1 \mathfrak{Y}_2}\right)^{-\frac{1}{2}} 2\mathfrak{Y}_1 \mathfrak{Y}_2 \lambda^2 \quad (3)$$

The variance of the statistic $\mathcal{T}_{\mathfrak{R}}$ is calculated as follows:

$$\begin{aligned} \text{Var}(\mathcal{T}_{\mathfrak{R}}(\mathfrak{S}_1, \mathfrak{S}_2)) &= \text{Var}\left(\left(\frac{\mathfrak{Y}_1 + \mathfrak{Y}_2}{\mathfrak{Y}_1 \mathfrak{Y}_2}\right)^{-\frac{1}{2}} \sum_{i=1}^{\mathfrak{Y}_1} \sum_{j=1}^{\mathfrak{Y}_2} (\mathcal{U}_i, \mathcal{X}_j)\right) \\ &= \frac{\mathfrak{Y}_1 \mathfrak{Y}_2}{\mathfrak{Y}_1 + \mathfrak{Y}_2} \text{Var}\left(\sum_{i=1}^{\mathfrak{Y}_1} \sum_{j=1}^{\mathfrak{Y}_2} (\mathcal{D}(\mathcal{U}_i, \mathcal{X}_j))\right) \end{aligned}$$

Since $\mathcal{D}(\mathcal{U}_i, \mathcal{X}_j) = (\mathcal{U}_i - \mathcal{X}_j)^2$, $i= 1, 2, \dots, \mathfrak{Y}_1$; $j= 1, 2, \dots, \mathfrak{Y}_2$, then we get:

$$\text{Var}(\mathcal{T}_{\mathfrak{R}}(\mathfrak{S}_1, \mathfrak{S}_2)) = \frac{\mathfrak{Y}_1 \mathfrak{Y}_2}{\mathfrak{Y}_1 + \mathfrak{Y}_2} \text{Var}\left(\sum_{i=1}^{\mathfrak{Y}_1} \sum_{j=1}^{\mathfrak{Y}_2} (\mathcal{U}_i - \mathcal{X}_j)^2\right) \quad (4)$$

The formula $\text{var}(\sum_{i=1}^{\mathfrak{Y}_1} \sum_{j=1}^{\mathfrak{Y}_2} (\mathcal{U}_i - \mathcal{X}_j)^2)$ Assuming independent and identically distributed (i.i.d.) normal samples, the following computation is made in Equation (4):by setting $\mathcal{D}_{ij} = (\mathcal{U}_i - \mathcal{X}_j)^2$ we get :

$$\begin{aligned} \text{Var}\left(\sum_{i=1}^{\mathfrak{Y}_1} \sum_{j=1}^{\mathfrak{Y}_2} (\mathcal{U}_i - \mathcal{X}_j)^2\right) &= [Var(\mathcal{D}_{11}) + Var(\mathcal{D}_{12}) + \dots + Var(\mathcal{D}_{1\mathfrak{Y}_2}) + Var(\mathcal{D}_{21}) \\ &+ Var(\mathcal{D}_{22}) + \dots + Var(\mathcal{D}_{2\mathfrak{Y}_2}) + \dots + Var(\mathcal{D}_{\mathfrak{Y}_1 1}) \\ &+ Var(\mathcal{D}_{\mathfrak{Y}_1 2}) + \dots + Var(\mathcal{D}_{\mathfrak{Y}_1 \mathfrak{Y}_2})] \\ &+ 2[cov(\mathcal{D}_{11}, \mathcal{D}_{12}) + cov(\mathcal{D}_{11}, \mathcal{D}_{13}) + \dots + cov(\mathcal{D}_{11}, \mathcal{D}_{1\mathfrak{Y}_2}) \\ &+ cov(\mathcal{D}_{12}, \mathcal{D}_{13}) + cov(\mathcal{D}_{12}, \mathcal{D}_{14}) + \dots + cov(\mathcal{D}_{12}, \mathcal{D}_{1\mathfrak{Y}_2}) \\ &+ \dots + cov(\mathcal{D}_{\mathfrak{Y}_1(\mathfrak{Y}_2-1)}, \mathcal{D}_{\mathfrak{Y}_1 \mathfrak{Y}_2})] \end{aligned} \quad (5)$$

since,

$$\begin{aligned}
 \text{Var}(\mathfrak{D}) &= \text{Var}((\mathfrak{U} - \mathfrak{X})^2) = E((\mathfrak{U} - \mathfrak{X})^4) - [E((\mathfrak{U} - \mathfrak{X})^2)]^2 \\
 &= E(\mathfrak{U}^4 - 4\mathfrak{U}^3\mathfrak{X} + 6\mathfrak{U}^2\mathfrak{X}^2 - 4\mathfrak{U}\mathfrak{X}^3 + \mathfrak{X}^4) - (E(\mathfrak{U}^2 - 2\mathfrak{U}\mathfrak{X} + \mathfrak{X}^2))^2 \\
 &= E(\mathfrak{U}^4) - 4E(\mathfrak{U}^3)E(\mathfrak{X}) + 6E(\mathfrak{U}^2)E(\mathfrak{X}^2) - 4E(\mathfrak{U})E(\mathfrak{X}^3) + E(\mathfrak{X}^4) \\
 &\quad - (E(\mathfrak{U}^2) - 2E(\mathfrak{U})E(\mathfrak{X}) + E(\mathfrak{X}^2))^2 \\
 &= 24\lambda^4 - 4(6\lambda^3)(\lambda) + 6(2\lambda^2)(2\lambda^2) - 4(\lambda)(6\lambda^3) + 24\lambda^4 \\
 &\quad - (2\lambda^2 - 2\lambda^2 + 2\lambda^2)^2 \\
 &= 24\lambda^4 - (2\lambda^2)^2 \\
 &= 20\lambda^4 \tag{6}
 \end{aligned}$$

since $\mathfrak{D}_{11} = (\mathfrak{U}_1 - \mathfrak{X}_1)^2$ and $\mathfrak{D}_{12} = (\mathfrak{U}_1 - \mathfrak{X}_2)^2$

$$\begin{aligned}
 \text{Cov}(\mathfrak{D}_{11}, \mathfrak{D}_{12}) &= E(\mathfrak{D}_{11}\mathfrak{D}_{12}) - E(\mathfrak{D}_{11})E(\mathfrak{D}_{12}) \\
 &= E((\mathfrak{U}_1 - \mathfrak{X}_1)^2(\mathfrak{U}_1 - \mathfrak{X}_2)^2) - E((\mathfrak{U}_1 - \mathfrak{X}_1)^2)E((\mathfrak{U}_1 - \mathfrak{X}_2)^2) \\
 &= E(\mathfrak{U}_1^4 - 2\mathfrak{U}_1^3\mathfrak{X}_2 + \mathfrak{U}_1^2\mathfrak{X}_2^2 - 2\mathfrak{U}_1^3\mathfrak{X}_1 + 4\mathfrak{U}_1^2\mathfrak{X}_1\mathfrak{X}_2 - 2\mathfrak{U}_1\mathfrak{X}_1\mathfrak{X}_2^2 + \mathfrak{U}_1^2\mathfrak{X}_2^2 \\
 &\quad - 2\mathfrak{U}_1\mathfrak{X}_1^2\mathfrak{X}_2 + \mathfrak{X}_1^2\mathfrak{X}_2^2) - E(\mathfrak{U}_1^2 - 2\mathfrak{U}_1\mathfrak{X}_1 + \mathfrak{X}_1^2)E(\mathfrak{U}_1^2 - 2\mathfrak{U}_1\mathfrak{X}_2 + \mathfrak{X}_2^2) \\
 &= 24\lambda^4 - 12\lambda^4 + 4\lambda^4 - 12\lambda^4 + 8\lambda^4 - 4\lambda^4 + 4\lambda^4 - 4\lambda^4 + 4\lambda^4 \\
 &\quad - (2\lambda^2 - 2\lambda^2 + 2\lambda^2)(2\lambda^2 - 2\lambda^2 + 2\lambda^2) \\
 &= 12\lambda^4 - 4\lambda^4 \\
 &= 8\lambda^4 \tag{7}
 \end{aligned}$$

$$\begin{aligned}
 E(\mathfrak{D}) &= E((\mathfrak{U} - \mathfrak{X})^2) = E(\mathfrak{U}^2 - 2\mathfrak{U}\mathfrak{X} + \mathfrak{X}^2) = E(\mathfrak{U}^2) - 2E(\mathfrak{U})E(\mathfrak{X}) + E(\mathfrak{X}^2) \\
 &= 2\lambda^2 - 2\lambda^2 + 2\lambda^2 = 2\lambda^2 \tag{8}
 \end{aligned}$$

Applying Equations (6), (7), and (8) in Equation (5) we get the following

$$\text{Var}\left(\sum_{i=1}^{\mathfrak{Y}_1} \sum_{j=1}^{\mathfrak{Y}_2} (\mathfrak{U}_i - \mathfrak{X}_j)^2\right) = [20\lambda^4 + 16(\mathfrak{Y}_1 + \mathfrak{Y}_2 - 2)\lambda^4] \tag{9}$$

Applying Equation (9) in Equation (4) we get the variance of the statistic $\mathcal{T}_{\mathfrak{R}}$ which is given by

$$\text{Var}(\mathcal{T}_{\mathfrak{R}}(\mathfrak{S}_1, \mathfrak{S}_2)) = \frac{(\mathfrak{Y}_1 \mathfrak{Y}_2)^2}{\mathfrak{Y}_1 + \mathfrak{Y}_2} [20\lambda^4 + 16(\mathfrak{Y}_1 + \mathfrak{Y}_2 - 2)\lambda^4] \quad (10)$$

2.2- The expected value for the statistic $\mathcal{T}_{\mathfrak{R}}$ is computed as follows: Let \mathfrak{U} and \mathfrak{X} be two random variables that are independent such that $\mathfrak{U} \sim \text{Be}(\alpha, \beta)$ and $\mathfrak{X} \sim \text{Be}(\alpha, \beta)$.

Let $\mathfrak{S}_1 = \{\mathfrak{u}_1, \mathfrak{u}_2, \dots, \mathfrak{u}_{\mathfrak{Y}_1}\}$ and $\mathfrak{S}_2 = \{\mathfrak{x}_1, \mathfrak{x}_2, \dots, \mathfrak{x}_{\mathfrak{Y}_2}\}$ be two random samples from these random variables.

$$\begin{aligned} E(\mathcal{T}_{\mathfrak{R}}(\mathfrak{S}_1, \mathfrak{S}_2)) &= E\left(\left(\frac{\mathfrak{Y}_1 + \mathfrak{Y}_2}{\mathfrak{Y}_1 \mathfrak{Y}_2}\right)^{-\frac{1}{2}} \sum_{i=1}^{\mathfrak{Y}_1} \sum_{j=1}^{\mathfrak{Y}_2} \mathfrak{D}(\mathfrak{u}_i, \mathfrak{x}_j)\right) \\ &= \left(\frac{\mathfrak{Y}_1 + \mathfrak{Y}_2}{\mathfrak{Y}_1 \mathfrak{Y}_2}\right)^{-\frac{1}{2}} \sum_{i=1}^{\mathfrak{Y}_1} \sum_{j=1}^{\mathfrak{Y}_2} E(\mathfrak{D}(\mathfrak{u}_i, \mathfrak{x}_j)) \end{aligned}$$

Since $\mathfrak{D}(\mathfrak{u}_i, \mathfrak{x}_j) = (\mathfrak{u}_i - \mathfrak{x}_j)^2$, $i = 1, 2, \dots, \mathfrak{Y}_1$; $j = 1, 2, \dots, \mathfrak{Y}_2$, then we get

$$E(\mathcal{T}_{\mathfrak{R}}(\mathfrak{S}_1, \mathfrak{S}_2)) = \left(\frac{\mathfrak{Y}_1 + \mathfrak{Y}_2}{\mathfrak{Y}_1 \mathfrak{Y}_2}\right)^{-\frac{1}{2}} \sum_{i=1}^{\mathfrak{Y}_1} \sum_{j=1}^{\mathfrak{Y}_2} E((\mathfrak{u}_i - \mathfrak{x}_j)^2) \quad (11)$$

Equation (11) formula $E[(\mathfrak{u}_i - \mathfrak{x}_j)^2]$ is calculated as follows, assuming independent and identically distributed (i.i.d.) normal samples:

$$E((\mathfrak{U} - \mathfrak{X})^2) = \int_0^{\infty} \int_0^{\infty} (\mathfrak{u} - \mathfrak{x})^2 f(\mathfrak{u}, \mathfrak{x}) d\mathfrak{u} d\mathfrak{x} = \int_0^{\infty} \int_0^{\infty} (\mathfrak{u}^2 - 2\mathfrak{u}\mathfrak{x} + \mathfrak{x}^2) f(\mathfrak{u}, \mathfrak{x}) d\mathfrak{u} d\mathfrak{x}$$

After performing some mathematical processes, we get the following because the random variables \mathfrak{U} and \mathfrak{X} are equivalently distributed and independent:

$$\begin{aligned} E((\mathfrak{U} - \mathfrak{X})^2) &= E(\mathfrak{U}^2) - 2E(\mathfrak{U})E(\mathfrak{X}) + E(\mathfrak{X}^2) \\ &= \alpha\beta^2 + \alpha^2\beta^2 - 2\alpha^2\beta^2 + \alpha\beta^2 + \alpha^2\beta^2 = 2\alpha\beta^2 \end{aligned}$$

So, Equation (11) will equal to

$$E(\mathcal{T}_{\mathfrak{R}}(\mathfrak{S}_1, \mathfrak{S}_2)) = \left(\frac{\mathfrak{Y}_1 + \mathfrak{Y}_2}{\mathfrak{Y}_1 \mathfrak{Y}_2}\right)^{-\frac{1}{2}} \sum_{i=1}^{\mathfrak{Y}_1} \sum_{j=1}^{\mathfrak{Y}_2} 2\alpha\beta^2 = \left(\frac{\mathfrak{Y}_1 + \mathfrak{Y}_2}{\mathfrak{Y}_1 \mathfrak{Y}_2}\right)^{-\frac{1}{2}} 2\mathfrak{Y}_1 \mathfrak{Y}_2 \alpha\beta^2 \quad (12)$$

The variance of the statistic $\mathcal{T}_{\mathfrak{R}}$ is calculated as follows:

$$\begin{aligned} \text{Var}(\mathcal{T}_{\mathfrak{R}}(\mathfrak{S}_1, \mathfrak{S}_2)) &= \text{Var}\left(\left(\frac{\mathfrak{Y}_1 + \mathfrak{Y}_2}{\mathfrak{Y}_1 \mathfrak{Y}_2}\right)^{-\frac{1}{2}} \sum_{i=1}^{\mathfrak{Y}_1} \sum_{j=1}^{\mathfrak{Y}_2} \mathfrak{D}(\mathfrak{U}_i, \mathfrak{T}_j)\right) \\ &= \frac{\mathfrak{Y}_1 \mathfrak{Y}_2}{\mathfrak{Y}_1 + \mathfrak{Y}_2} \text{Var}\left(\sum_{i=1}^{\mathfrak{Y}_1} \sum_{j=1}^{\mathfrak{Y}_2} \mathfrak{D}(\mathfrak{U}_i, \mathfrak{T}_j)\right) \end{aligned}$$

Since $\mathfrak{D}(\mathfrak{U}_i, \mathfrak{T}_j) = (\mathfrak{U}_i - \mathfrak{T}_j)^2$, $i=1, 2, \dots, \mathfrak{Y}_1$; $j=1, 2, \dots, \mathfrak{Y}_2$, then we get

$$\text{Var}(\mathcal{T}_{\mathfrak{R}}(\mathfrak{S}_1, \mathfrak{S}_2)) = \frac{\mathfrak{Y}_1 \mathfrak{Y}_2}{\mathfrak{Y}_1 + \mathfrak{Y}_2} \text{Var}\left(\sum_{i=1}^{\mathfrak{Y}_1} \sum_{j=1}^{\mathfrak{Y}_2} (\mathfrak{U}_i - \mathfrak{T}_j)^2\right) \quad (13)$$

Equation (13) formula $\text{Var}\left(\sum_{i=1}^{\mathfrak{Y}_1} \sum_{j=1}^{\mathfrak{Y}_2} (\mathfrak{U}_i - \mathfrak{T}_j)^2\right)$ is calculated using the following method assuming (i.i.d) normal samples:

by setting $\mathfrak{D}_{ij} = (\mathfrak{U}_i - \mathfrak{T}_j)^2$ we get :

$$\begin{aligned} \text{Var}\left(\sum_{i=1}^{\mathfrak{Y}_1} \sum_{j=1}^{\mathfrak{Y}_2} (\mathfrak{U}_i - \mathfrak{T}_j)^2\right) &= [\text{Var}(\mathfrak{D}_{11}) + \text{Var}(\mathfrak{D}_{12}) + \dots + \text{Var}(\mathfrak{D}_{1\mathfrak{Y}_2}) + \text{Var}(\mathfrak{D}_{21}) \\ &+ \text{Var}(\mathfrak{D}_{22}) + \dots + \text{Var}(\mathfrak{D}_{2\mathfrak{Y}_2}) + \dots + \text{Var}(\mathfrak{D}_{\mathfrak{Y}_1 1}) \\ &+ \text{Var}(\mathfrak{D}_{\mathfrak{Y}_1 2}) + \dots + \text{Var}(\mathfrak{D}_{\mathfrak{Y}_1 \mathfrak{Y}_2})] \\ &+ 2[\text{cov}(\mathfrak{D}_{11}, \mathfrak{D}_{12}) + \text{cov}(\mathfrak{D}_{11}, \mathfrak{D}_{13}) + \dots + \text{cov}(\mathfrak{D}_{11}, \mathfrak{D}_{1\mathfrak{Y}_2}) \\ &+ \text{cov}(\mathfrak{D}_{12}, \mathfrak{D}_{13}) + \text{cov}(\mathfrak{D}_{12}, \mathfrak{D}_{14}) + \dots + \text{cov}(\mathfrak{D}_{12}, \mathfrak{D}_{1\mathfrak{Y}_2}) \\ &+ \dots + \text{cov}(\mathfrak{D}_{\mathfrak{Y}_1(\mathfrak{Y}_2-1)}, \mathfrak{D}_{\mathfrak{Y}_1 \mathfrak{Y}_2})] \end{aligned} \quad (14)$$

since,

$$\begin{aligned} \text{Var}(\mathfrak{D}) &= \text{Var}((\mathfrak{U} - \mathfrak{T})^2) = \text{E}((\mathfrak{U} - \mathfrak{T})^4) - [\text{E}((\mathfrak{U} - \mathfrak{T})^2)]^2 \\ &= \text{E}(\mathfrak{U}^4 - 4\mathfrak{U}^3\mathfrak{T} + 6\mathfrak{U}^2\mathfrak{T}^2 - 4\mathfrak{U}\mathfrak{T}^3 + \mathfrak{T}^4) - (\text{E}(\mathfrak{U}^2 - 2\mathfrak{U}\mathfrak{T} + \mathfrak{T}^2))^2 \\ &= \text{E}(\mathfrak{U}^4) - 4\text{E}(\mathfrak{U}^3)\text{E}(\mathfrak{T}) + 6\text{E}(\mathfrak{U}^2)\text{E}(\mathfrak{T}^2) - 4\text{E}(\mathfrak{U})\text{E}(\mathfrak{T}^3) + \text{E}(\mathfrak{T}^4) \\ &\quad - (\text{E}(\mathfrak{U}^2) - 2\text{E}(\mathfrak{U})\text{E}(\mathfrak{T}) + \text{E}(\mathfrak{T}^2))^2 \\ &= (\alpha^4\beta^4 + 6\alpha^3\beta^4 + 11\alpha^2\beta^4 + 6\alpha\beta^4) - 4(\alpha\beta)(\alpha^3\beta^3 + 3\alpha^2\beta^3 + 2\alpha\beta^3) \\ &\quad + 6(\alpha^2\beta^2 + \alpha\beta^2)^2 - 4(\alpha\beta)(\alpha^3\beta^3 + 3\alpha^2\beta^3 + 2\alpha\beta^3) \\ &\quad + (\alpha^4\beta^4 + 6\alpha^3\beta^4 + 11\alpha^2\beta^4 + 6\alpha\beta^4) \\ &\quad - (\alpha\beta^2 + \alpha^2\beta^2 - 2\alpha^2\beta^2 + \alpha\beta^2 + \alpha^2\beta^2)^2 \\ &= 2(\alpha^4\beta^4 + 6\alpha^3\beta^4 + 11\alpha^2\beta^4 + 6\alpha\beta^4) - 8(\alpha\beta)(\alpha^3\beta^3 + 3\alpha^2\beta^3 + 2\alpha\beta^3) \\ &\quad + 6(\alpha^2\beta^2 + \alpha\beta^2)^2 - (2\alpha\beta^2)^2 \\ &= 12\alpha^2\beta^4 + 12\alpha\beta^4 - 4\alpha^2\beta^4 \end{aligned}$$

$$= 8\alpha^2\beta^4 + 12\alpha\beta^4 \tag{15}$$

$$\begin{aligned} \text{Cov}(\mathcal{D}_{11}, \mathcal{D}_{12}) &= E(\mathcal{D}_{11}\mathcal{D}_{12}) - E(\mathcal{D}_{11})E(\mathcal{D}_{12}) \\ &= E((\mathcal{U}_1 - \mathcal{I}_1)^2(\mathcal{U}_1 - \mathcal{I}_2)^2) - E((\mathcal{U}_1 - \mathcal{I}_1)^2)E((\mathcal{U}_1 - \mathcal{I}_2)^2) \\ &= E(\mathcal{U}_1^4 - 2\mathcal{U}_1^3\mathcal{I}_2 + \mathcal{U}_1^2\mathcal{I}_2^2 - 2\mathcal{U}_1^3\mathcal{I}_1 + 4\mathcal{U}_1^2\mathcal{I}_1\mathcal{I}_2 - 2\mathcal{U}_1\mathcal{I}_1\mathcal{I}_2^2 + \mathcal{U}_1^2\mathcal{I}_2^2 \\ &\quad - 2\mathcal{U}_1\mathcal{I}_1^2\mathcal{I}_2 + \mathcal{I}_1^2\mathcal{I}_2^2) - E(\mathcal{U}_1^2 - 2\mathcal{U}_1\mathcal{I}_1 + \mathcal{I}_1^2)E(\mathcal{U}_1^2 - 2\mathcal{U}_1\mathcal{I}_2 + \mathcal{I}_2^2) \\ &= (\alpha^4\beta^4 + 6\alpha^3\beta^4 + 11\alpha^2\beta^4 + 6\alpha\beta^4) - 2(\alpha\beta)(\alpha^3\beta^3 + 3\alpha^2\beta^3 + 2\alpha\beta^3) \\ &\quad + (\alpha^2\beta^2 + \alpha\beta^2)^2 - 2(\alpha\beta)(\alpha^3\beta^3 + 3\alpha^2\beta^3 + 2\alpha\beta^3) \\ &\quad + 4(\alpha^2\beta^2 + \alpha\beta^2)(\alpha^2\beta^2) - 2(\alpha^2\beta^2 + \alpha\beta^2)(\alpha^2\beta^2) + (\alpha^2\beta^2 + \alpha\beta^2)^2 \\ &\quad - 2(\alpha^2\beta^2 + \alpha\beta^2)(\alpha^2\beta^2) + (\alpha^2\beta^2 + \alpha\beta^2)^2 - (\alpha\beta^2 + \alpha^2\beta^2 - 2\alpha^2\beta^2 \\ &\quad + \alpha\beta^2 + \alpha^2\beta^2)(\alpha\beta^2 + \alpha^2\beta^2 - 2\alpha^2\beta^2 + \alpha\beta^2 + \alpha^2\beta^2) \\ &= (\alpha^4\beta^4 + 6\alpha^3\beta^4 + 11\alpha^2\beta^4 + 6\alpha\beta^4) - 4(\alpha\beta)(\alpha^3\beta^3 + 3\alpha^2\beta^3 + 2\alpha\beta^3) \\ &\quad + 3(\alpha^2\beta^2 + \alpha\beta^2)^2 - (2\alpha\beta^2)^2 \\ &= 6\alpha^2\beta^4 + 6\alpha\beta^4 - 4\alpha^2\beta^4 \end{aligned}$$

$$= 2\alpha^2\beta^4 + 6\alpha\beta^4 \tag{16}$$

$$\begin{aligned} E(\mathcal{D}) &= E((\mathcal{U} - \mathcal{I})^2) = E(\mathcal{U}^2 - 2\mathcal{U}\mathcal{I} + \mathcal{I}^2) = E(\mathcal{U}^2) - 2E(\mathcal{U})E(\mathcal{I}) + E(\mathcal{I}^2) \\ &= \alpha\beta^2 + \alpha^2\beta^2 - 2\alpha^2\beta^2 + \alpha\beta^2 + \alpha^2\beta^2 = 2\alpha\beta^2 \end{aligned}$$

Applying Equations (15), (16), and (17) in Equation (14) we get the following

$$\begin{aligned} &\text{Var}\left(\sum_{i=1}^{\mathcal{Y}_1} \sum_{j=1}^{\mathcal{Y}_2} (\mathcal{U}_i - \mathcal{I}_j)^2\right) \\ &= [8\alpha^2\beta^4 + 12\alpha\beta^4 + (\mathcal{Y}_1 + \mathcal{Y}_2 - 2)(4\alpha^2\beta^4 + 12\alpha\beta^4)] \end{aligned} \tag{17}$$

Applying Equation (17) in Equation (13) we get the variance of the statistic $\mathcal{T}_{\mathfrak{R}}$ which is given by

$$\begin{aligned} \text{Var}(\mathcal{T}_{\mathfrak{R}}(\mathfrak{S}_1, \mathfrak{S}_2)) &= \frac{(\mathfrak{Y}_1 \mathfrak{Y}_2)^2}{\mathfrak{Y}_1 + \mathfrak{Y}_2} [8\alpha^2 \beta^4 + 12\alpha \beta^4] \\ &+ (\mathfrak{Y}_1 + \mathfrak{Y}_2 - 2)(4\alpha^2 \beta^4 + 12\alpha \beta^4) \end{aligned} \tag{18}$$

3. PROPERTIES OF THE ENERGY-LIKE STATISTIC WITH P-DIMENSIONAL CASE ($p > 1$)

3.1-Let \mathfrak{U} and \mathfrak{X} in R^p , $p > 1$ be two independent random variables such that $\mathfrak{U} \sim \text{Exp}(\lambda)$ and $\mathfrak{X} \sim \text{Exp}(\lambda)$, and suppose \mathfrak{S}_1 and \mathfrak{S}_2 be independent random samples of \mathfrak{U} and \mathfrak{X} , respectively, which are give by $\mathfrak{S}_1 = \{(\mathfrak{U}_{11}, \mathfrak{U}_{21}, \dots, \mathfrak{U}_{p1}), (\mathfrak{U}_{12}, \mathfrak{U}_{22}, \dots, \mathfrak{U}_{p2}), \dots, (\mathfrak{U}_{1\mathfrak{Y}_1}, \mathfrak{U}_{2\mathfrak{Y}_1}, \dots, \mathfrak{U}_{p\mathfrak{Y}_1})\}$ and $\mathfrak{S}_2 = \{(\mathfrak{X}_{11}, \mathfrak{X}_{21}, \dots, \mathfrak{X}_{p1}), (\mathfrak{X}_{12}, \mathfrak{X}_{22}, \dots, \mathfrak{X}_{p2}), \dots, (\mathfrak{X}_{1\mathfrak{Y}_2}, \mathfrak{X}_{2\mathfrak{Y}_2}, \dots, \mathfrak{X}_{p\mathfrak{Y}_2})\}$. For the high-dimensional case, the expected value and the variance of the $\mathcal{T}_{\mathfrak{R}}$ statistic under squared Euclidean distance measure are defined by:

$$\begin{aligned} E(\mathcal{T}_{\mathfrak{R}}(\mathfrak{S}_1, \mathfrak{S}_2)) &= E \left(\left(\frac{\mathfrak{Y}_1 + \mathfrak{Y}_2}{\mathfrak{Y}_1 \mathfrak{Y}_2} \right)^{-\frac{1}{2}} \sum_{i=1}^{\mathfrak{Y}_1} \sum_{j=1}^{\mathfrak{Y}_2} (\mathfrak{D}(\mathfrak{U}_{1i}, \mathfrak{X}_{1j}) + \mathfrak{D}(\mathfrak{U}_{2i}, \mathfrak{X}_{2j}) + \dots \right. \\ &\quad \left. + \mathfrak{D}(\mathfrak{U}_{\mathfrak{Y}_1 i}, \mathfrak{X}_{\mathfrak{Y}_2 j})) \right) \\ &= \left(\frac{\mathfrak{Y}_1 + \mathfrak{Y}_2}{\mathfrak{Y}_1 \mathfrak{Y}_2} \right)^{-\frac{1}{2}} \left(\sum_{i=1}^{\mathfrak{Y}_1} \sum_{j=1}^{\mathfrak{Y}_2} E(\mathfrak{D}(\mathfrak{U}_{1i}, \mathfrak{X}_{1j}) + \mathfrak{D}(\mathfrak{U}_{2i}, \mathfrak{X}_{2j}) + \dots \right. \\ &\quad \left. + \mathfrak{D}(\mathfrak{U}_{\mathfrak{Y}_1 i}, \mathfrak{X}_{\mathfrak{Y}_2 j})) \right) \end{aligned}$$

Since $\mathfrak{D}(\mathfrak{U}, \mathfrak{X}) = (\mathfrak{U} - \mathfrak{X})^2$, then we get

$$\begin{aligned} E(\mathcal{T}_{\mathfrak{R}}(\mathfrak{S}_1, \mathfrak{S}_2)) &= \left(\frac{\mathfrak{Y}_1 + \mathfrak{Y}_2}{\mathfrak{Y}_1 \mathfrak{Y}_2} \right)^{-\frac{1}{2}} \left(\sum_{i=1}^{\mathfrak{Y}_1} \sum_{j=1}^{\mathfrak{Y}_2} E((\mathfrak{U}_{1i} - \mathfrak{X}_{1j})^2 + (\mathfrak{U}_{2i} - \mathfrak{X}_{2j})^2 + \dots \right. \\ &\quad \left. + (\mathfrak{U}_{\mathfrak{Y}_1 i} - \mathfrak{X}_{\mathfrak{Y}_2 j})^2) \right) \end{aligned} \tag{19}$$

Since $E(\mathfrak{U} - \mathfrak{X})^2 = 2\lambda^2$, then we get

$$E(\mathcal{T}_{\mathfrak{R}}(\mathfrak{S}_1, \mathfrak{S}_2)) = \left(\frac{\mathfrak{Y}_1 + \mathfrak{Y}_2}{\mathfrak{Y}_1 \mathfrak{Y}_2} \right)^{-\frac{1}{2}} 2p\mathfrak{Y}_1 \mathfrak{Y}_2 \lambda^2 \tag{20}$$

The variance of the statistic $\mathcal{T}_{\mathfrak{R}}(\mathfrak{S}_1, \mathfrak{S}_2)$ is calculated as follows:

$$\begin{aligned} \text{Var}(\mathcal{T}_{\mathbb{R}}(\mathfrak{S}_1, \mathfrak{S}_2)) &= \text{Var} \left(\left(\frac{\mathfrak{Y}_1 + \mathfrak{Y}_2}{\mathfrak{Y}_1 \mathfrak{Y}_2} \right)^{-\frac{1}{2}} \sum_{i=1}^{\mathfrak{Y}_1} \sum_{j=1}^{\mathfrak{Y}_2} (\mathfrak{D}(\mathfrak{U}_{1i}, \mathfrak{I}_{1j}) + \mathfrak{D}(\mathfrak{U}_{2i}, \mathfrak{I}_{2j}) + \dots \right. \\ &\quad \left. + \mathfrak{D}(\mathfrak{U}_{\mathfrak{Y}_1 i}, \mathfrak{I}_{\mathfrak{Y}_2 j}) \right) \\ &= \left(\frac{\mathfrak{Y}_1 \mathfrak{Y}_2}{\mathfrak{Y}_1 + \mathfrak{Y}_2} \right) \text{Var} \left(\sum_{i=1}^{\mathfrak{Y}_1} \sum_{j=1}^{\mathfrak{Y}_2} (\mathfrak{D}(\mathfrak{U}_{1i}, \mathfrak{I}_{1j}) + \mathfrak{D}(\mathfrak{U}_{2i}, \mathfrak{I}_{2j}) + \dots \right. \\ &\quad \left. + \mathfrak{D}(\mathfrak{U}_{\mathfrak{Y}_1 i}, \mathfrak{I}_{\mathfrak{Y}_2 j}) \right) \end{aligned}$$

Since $\mathfrak{D}(\mathfrak{U}, \mathfrak{I}) = (\mathfrak{U} - \mathfrak{I})^2$, then we get

$$\begin{aligned} \text{Var}(\mathcal{T}_{\mathbb{R}}(\mathfrak{S}_1, \mathfrak{S}_2)) &= \left(\frac{\mathfrak{Y}_1 \mathfrak{Y}_2}{\mathfrak{Y}_1 + \mathfrak{Y}_2} \right) \text{Var} \left(\sum_{i=1}^{\mathfrak{Y}_1} \sum_{j=1}^{\mathfrak{Y}_2} [(\mathfrak{U}_{1i} - \mathfrak{I}_{1j})^2 + (\mathfrak{U}_{2i} - \mathfrak{I}_{2j})^2 + \dots \right. \\ &\quad \left. + (\mathfrak{U}_{\mathfrak{Y}_1 i} - \mathfrak{I}_{\mathfrak{Y}_2 j})^2] \right) \end{aligned} \tag{21}$$

Applying Equation (9) in Equation (21) we get the variance of the statistic $\mathcal{T}_{\mathbb{R}}$ which is given by

$$\text{Var}(\mathcal{T}_{\mathbb{R}}(\mathfrak{S}_1, \mathfrak{S}_2)) = \frac{p(\mathfrak{Y}_1 \mathfrak{Y}_2)^2}{\mathfrak{Y}_1 + \mathfrak{Y}_2} [20\lambda^4 + 16(\mathfrak{Y}_1 + \mathfrak{Y}_2 - 2)\lambda^4] \tag{22}$$

3.2-Let \mathfrak{U} and \mathfrak{I} in \mathbb{R}^p , $p > 1$ be two independent random variables such that $\mathfrak{U} \sim \text{Be}(\alpha, \beta)$ and $\mathfrak{I} \sim \text{Be}(\alpha, \beta)$. Assume that \mathfrak{S}_1 and \mathfrak{S}_2 are independent random samples of \mathfrak{U} and \mathfrak{I} , respectively, given by $\mathfrak{S}_1 = \{(\mathfrak{U}_{11}, \mathfrak{U}_{21}, \dots, \mathfrak{U}_{p1}), (\mathfrak{U}_{12}, \mathfrak{U}_{22}, \dots, \mathfrak{U}_{p2}), \dots, (\mathfrak{U}_{1\mathfrak{Y}_1}, \mathfrak{U}_{2\mathfrak{Y}_1}, \dots, \mathfrak{U}_{p\mathfrak{Y}_1})\}$ and $\mathfrak{S}_2 = \{(\mathfrak{I}_{11}, \mathfrak{I}_{21}, \dots, \mathfrak{I}_{p1}), (\mathfrak{I}_{12}, \mathfrak{I}_{22}, \dots, \mathfrak{I}_{p2}), \dots, (\mathfrak{I}_{1\mathfrak{Y}_2}, \mathfrak{I}_{2\mathfrak{Y}_2}, \dots, \mathfrak{I}_{p\mathfrak{Y}_2})\}$. For the high-dimensional situation, the expected value and the variance of the $\mathcal{T}_{\mathbb{R}}$ statistic under squared Euclidean distance measure are determined by:

$$\begin{aligned}
 E(\mathcal{T}_{\bar{R}}(\mathcal{G}_1, \mathcal{G}_2)) &= E\left(\left(\frac{\mathfrak{Y}_1 + \mathfrak{Y}_2}{\mathfrak{Y}_1 \mathfrak{Y}_2}\right)^{-\frac{1}{2}} \sum_{i=1}^{\mathfrak{Y}_1} \sum_{j=1}^{\mathfrak{Y}_2} (\mathfrak{D}(u_{1i}, \mathfrak{T}_{1j}) + \mathfrak{D}(u_{2i}, \mathfrak{T}_{2j}) + \dots \right. \\
 &\quad \left. + \mathfrak{D}(u_{\mathfrak{Y}_1 i}, \mathfrak{T}_{\mathfrak{Y}_2 j}))\right) \\
 &= \left(\frac{\mathfrak{Y}_1 + \mathfrak{Y}_2}{\mathfrak{Y}_1 \mathfrak{Y}_2}\right)^{-\frac{1}{2}} \left(\sum_{i=1}^{\mathfrak{Y}_1} \sum_{j=1}^{\mathfrak{Y}_2} E(\mathfrak{D}(u_{1i}, \mathfrak{T}_{1j}) + \mathfrak{D}(u_{2i}, \mathfrak{T}_{2j}) + \dots \right. \\
 &\quad \left. + \mathfrak{D}(u_{\mathfrak{Y}_1 i}, \mathfrak{T}_{\mathfrak{Y}_2 j}))\right)
 \end{aligned}$$

Since $\mathfrak{D}(u, \mathfrak{T}) = (u - \mathfrak{T})^2$, then we get

$$\begin{aligned}
 E(\mathcal{T}_{\bar{R}}(\mathcal{G}_1, \mathcal{G}_2)) &= \left(\frac{\mathfrak{Y}_1 + \mathfrak{Y}_2}{\mathfrak{Y}_1 \mathfrak{Y}_2}\right)^{-\frac{1}{2}} \left(\sum_{i=1}^{\mathfrak{Y}_1} \sum_{j=1}^{\mathfrak{Y}_2} E((u_{1i} - \mathfrak{T}_{1j})^2 + (u_{2i} - \mathfrak{T}_{2j})^2 + \dots \right. \\
 &\quad \left. + (u_{\mathfrak{Y}_1 i} - \mathfrak{T}_{\mathfrak{Y}_2 j})^2)\right) \tag{23}
 \end{aligned}$$

Since $E(u - \mathfrak{T})^2 = 2\alpha\beta^2$, then we get

$$E(\mathcal{T}_{\bar{R}}(\mathcal{G}_1, \mathcal{G}_2)) = \left(\frac{\mathfrak{Y}_1 + \mathfrak{Y}_2}{\mathfrak{Y}_1 \mathfrak{Y}_2}\right)^{-\frac{1}{2}} 2p\mathfrak{Y}_1 \mathfrak{Y}_2 \alpha\beta^2 \tag{24}$$

The variance of the statistic $\mathcal{T}_{\bar{R}}(\mathcal{G}_1, \mathcal{G}_2)$ is calculated as follows:

$$\begin{aligned}
 \text{Var}(\mathcal{T}_{\bar{R}}(\mathcal{G}_1, \mathcal{G}_2)) &= \text{Var}\left(\left(\frac{\mathfrak{Y}_1 + \mathfrak{Y}_2}{\mathfrak{Y}_1 \mathfrak{Y}_2}\right)^{-\frac{1}{2}} \sum_{i=1}^{\mathfrak{Y}_1} \sum_{j=1}^{\mathfrak{Y}_2} (\mathfrak{D}(u_{1i}, \mathfrak{T}_{1j}) + \mathfrak{D}(u_{2i}, \mathfrak{T}_{2j}) + \dots \right. \\
 &\quad \left. + \mathfrak{D}(u_{\mathfrak{Y}_1 i}, \mathfrak{T}_{\mathfrak{Y}_2 j}))\right) \\
 &= \left(\frac{\mathfrak{Y}_1 \mathfrak{Y}_2}{\mathfrak{Y}_1 + \mathfrak{Y}_2}\right) \text{Var}\left(\sum_{i=1}^{\mathfrak{Y}_1} \sum_{j=1}^{\mathfrak{Y}_2} (\mathfrak{D}(u_{1i}, \mathfrak{T}_{1j}) + \mathfrak{D}(u_{2i}, \mathfrak{T}_{2j}) + \dots \right. \\
 &\quad \left. + \mathfrak{D}(u_{\mathfrak{Y}_1 i}, \mathfrak{T}_{\mathfrak{Y}_2 j}))\right)
 \end{aligned}$$

Since $\mathfrak{D}(u, \mathfrak{T}) = (u - \mathfrak{T})^2$, then we get

$$\text{Var}(\mathcal{T}_{\mathfrak{R}}(\mathfrak{S}_1, \mathfrak{S}_2)) = \left(\frac{\mathfrak{Y}_1 \mathfrak{Y}_2}{\mathfrak{Y}_1 + \mathfrak{Y}_2} \right) \text{Var} \left(\sum_{i=1}^{\mathfrak{Y}_1} \sum_{j=1}^{\mathfrak{Y}_2} \left[(\mathfrak{u}_{1i} - \mathfrak{T}_{1j})^2 + (\mathfrak{u}_{2i} - \mathfrak{T}_{2j})^2 + \dots + (\mathfrak{u}_{\mathfrak{Y}_1 i} - \mathfrak{T}_{\mathfrak{Y}_2 j})^2 \right] \right) \quad (25)$$

Applying Equation (18) in Equation (25) we get the variance of the statistic $\mathcal{T}_{\mathfrak{R}}$ which is given by

$$\text{Var}(\mathcal{T}_{\mathfrak{R}}(\mathfrak{S}_1, \mathfrak{S}_2)) = \frac{p(\mathfrak{Y}_1 \mathfrak{Y}_2)^2}{\mathfrak{Y}_1 + \mathfrak{Y}_2} [8\alpha^2 \beta^4 + 12\alpha \beta^4 + (\mathfrak{Y}_1 + \mathfrak{Y}_2 - 2)(4\alpha^2 \beta^4 + 12\alpha \beta^4)] \quad (26)$$

4. CONCLUSION

This work describes the properties of an energy-like statistic based on distance in both one-dimensional and p-dimensional cases. The statistical test might be utilized as a two-sample test to find homogeneity between the distributions of two groups based on typical linear data.

REFERENCES

1. Little, T. D. (Ed.). (2013). *The Oxford handbook of quantitative methods, volume 1: Foundations*. Oxford University Press.
2. Rizzo, M. L. (2019). *Statistical computing with R*. CRC Press.
3. Ali, A. J., & Abushilah, S. F. (2022). Distribution-free two-sample homogeneity test for circular data based on geodesic distance. *International Journal of Nonlinear Analysis and Applications*, **13**(1), 2703-2711.
4. Rizzo, M. L., & Székely, G. J. (2016). Energy distance. *wiley interdisciplinary reviews: Computational statistics*, **8**(1), 27-38.
5. Ali, A. J., & Abushilah, S. F. (2023, May). Properties of novel energy-like statistic under normal samples and squared Euclidean distance. In *AIP Conference Proceedings* (Vol. **2593**, No. 1). AIP Publishing.