Nonparametric two sample tests for scale parameters of multivariate distributions

Atul R Chavan[1,a], Digambar T Shirke[a]

[a]Department of Statistics, Shivaji University, India

Abstract

In this paper, a notion of data depth is used to propose nonparametric multivariate two sample tests for difference between scale parameters. Data depth can be used to measure the centrality or outlying-ness of the multivariate data point relative to data cloud. A difference in the scale parameters indicates the difference in the depth values of a multivariate data point. By observing this fact on a depth vs depth plot (DD-plot), we propose nonparametric multivariate two sample tests for scale parameters of multivariate distributions. The p-values of these proposed tests are obtained by using Fisher's permutation approach. The power performance of these proposed tests has been reported for few symmetric and skewed multivariate distributions with the existing tests. Illustration with real-life data is also provided.

Keywords: data depth, scale parameters, two sample test, permutation test

1. Introduction

The problem of testing scale parameters of two multivariate distributions is well addressed in the literature. The popular parametric two sample Box M-test (Box, 1949) and F_{product} test (Bibby *et al.*, 1979) are used if both samples come from multivariate normal distribution. In the case of non-normality, alternative tests are used that are mostly nonparametric. Many tests have been proposed for testing scale difference between two multivariate distributions. Some notions of data depth are widely and currently used in statistical testing procedures. Data depth is a device that measures the centrality or outlying-ness of a multivariate data point with respect to a multivariate distribution or a data cloud. It produces a natural center-outward ranking of points. Many nonparametric two sample tests for location and/or scale parameters of multivariate distributions are developed with the help of such ranking. Liu and Singh (1993), Rousson (2002), Li *et al.* (2011), and Chavan and Shirke (2019) have proposed tests for testing location as well as scale parameters of two multivariate distributions. Many multivariate two sample nonparametric tests for location parameters are proposed by Li and Liu (2004), Chavan and Shirke (2016), Shirke and Khorate (2017), and Pawar and Shirke (2019). Chenouri (2004), Liu and Singh (2006), Chenouri and Small (2012), and Li and Liu (2016) also developed multivariate two sample scale tests based on data depth.

In the present paper, we propose multivariate nonparametric two sample tests for scale parameters using the notion of data depth. The centrality or outlying-ness of a multivariate data point is measured by the data depth function. Difference in scale parameters exhibits a difference in the depth values. By observing this phenomenon, two sample tests are proposed in this paper. Fisher's permutation

approach is used for obtaining the *p*-values of the proposed tests. The performance of the proposed tests has been evaluated in terms of the power and is compared it to the existing F_{product} test and depth based rank test.

The remaining part of the paper is ordered as follows. The notion of data depth and depth-versus-depth plot (DD-plot) is described in Section 2. Tests proposed by Liu and Singh (2006) and Bibby *et al.* (1979) for scale parameters of multivariate distributions are described in Section 3. The proposed two sample tests for scale parameters are discussed in Section 4. Section 5 is devoted to the performance study of the proposed tests and the applicability of the proposed test for real-life data is provided in Section 6. Conclusions are given in Section 7.

2. Notion of data depth and DD-plot

2.1. Data depth

Let $x_0 \in \mathbb{R}^p$ be any point then its depth with respect to the multivariate distribution F is denoted by $D(x_0, F)$. Tukey first coined the word depth in 1975. Data depth is a mapping from *p*-dimensional real space to an interval $[0, \infty)$. It measures the centrality or outlying-ness of a data point with respect to a data cloud. Data depth gives a natural center-outward ranking of a multivariate data points with respect to a data cloud that is utilized in various statistical methods. It includes the development of some nonparametric tests, construction of nonparametric control charts, nonparametric classification techniques, and outlier detection. In the context of data depth, the deepest point or points (point or points having maximum depth) is treated as a sample median.

The current literature show several depth functions. Some of the depth functions are simplicial depth (Liu, 1990), Mahalanobis depth (Mahalanobis, 1936), halfspace depth (Tukey, 1975), spatial depth (Serfling, 2002), Oja depth (Oja, 1983), majority depth (Singh, 1991), and likelihood depth (Fraiman *et al.*, 1999). One can refer Liu *et al.* (1999) for detailed information. According to Zuo and Serfling (2000), depth function should have properties such as maximality at a center, affine invariance, vanishing at infinity and monotonicity relative to deepest point. We now review four of these depth functions used in our simulation study.

- Simplicial depth

  Simplicial depth (SD(·)) of a point $x_0 \in \mathbb{R}^p$ relative to any multivariate distribution F on $\mathbb{R}^p$ is defined as,

  $$\text{SD}(x_0, F) = \Pr_F(s[X_1, X_2, \ldots, X_{p+1}] \ni x_0),$$

  where $X_1, X_2, \ldots, X_{p+1}$ are independent and identically distributed observations from F and $s[X_1, X_2, \ldots, X_{p+1}]$ is a closed simplex whose vertices are $X_1, X_2, \ldots, X_{p+1}$. The sample version of simplicial depth is,

  $$\text{SD}(x_0, F_m) = \binom{m}{p+1}^{-1} \sum_* I(x_0 \epsilon s[X_{i1}, X_{i2}, \ldots, X_{ip+1}]),$$

  where F_m is an empirical distribution function, (∗) runs over all possible subsets of size $(p + 1)$ of the sample $X_1, X_2, \ldots, X_m$ and $I(\cdot)$ is an usual indicator function. Larger the depth $\text{SD}(x_0, F_m)$ indicates x_0 is contained in more simplices generated from the sample.

- Mahalanobis depth

Mahalanobis depth (MD($\cdot$)) of a point $x_0 \in \mathbb{R}^p$ relative to any multivariate distribution F on $\mathbb{R}^p$ is defined as,

$$MD(x_0, F) = \left[1 + (x_0 - \theta)'\Sigma^{-1}(x_0 - \theta)\right]^{-1},$$

where θ and Σ are the location parameter or center and variance covariance matrix or dispersion matrix of F respectively. The sample version of Mahalanobis depth can be obtained by replacing θ by $\bar{X}$ (sample mean vector) and Σ by S (sample variance covariance matrix).

- Halfspace depth

Halfspace depth (HSD($\cdot$)) of a point $x_0 \in \mathbb{R}^p$ relative to any multivariate distribution F on $\mathbb{R}^p$ is defined as,

$$HSD(x_0, F) = \inf_H\{\Pr(H) : H \text{ is a closed halfspace which contains the point } x_0\}.$$

The sample version of HSD(x_0, F) is obtained by replacing F by F_m.

- Spatial depth

Spatial depth (SPD($\cdot$)) of a point $x_0 \in \mathbb{R}^p$ relative to any multivariate distribution F on $\mathbb{R}^p$ is defined as,

$$SPD(x_0, F) = 1 - \left\| \int S(x_0 - X)dF(X) \right\| = 1 - \|E(S(x_0 - X))\|,$$

where $\| \cdot \|$ is an usual Euclidean norm in $\mathbb{R}^p$ and $S : \mathbb{R}^p \rightarrow \mathbb{R}^p$ is the multivariate spatial sign function given by,

$$S(x_0) = \begin{cases} \dfrac{x_0}{\|x_0\|}, & \text{if } x_0 \neq 0, \\ 0, & \text{if } x_0 = 0. \end{cases}$$

The sample version of SPD(x_0, F) can be obtained by,

$$SPD(x_0, F_m) = 1 - \left\| \frac{1}{m} \sum_{i=1}^{m} S(x_0 - X_i) \right\|.$$

2.2. DD-plot

Let $\mathcal{X} = \{X_1, X_2, \ldots, X_m\}$ and $\mathcal{Y} = \{Y_1, Y_2, \ldots, Y_n\}$ be a two data vectors of sizes m and n observed from a two continuous multivariate distributions F and G respectively, where each $X_i \in \mathbb{R}^p$ and $Y_j \in \mathbb{R}^p$, $i = 1, 2, \ldots, m$, $j = 1, 2, \ldots, n$. Let $D(z, F_m)$ and $D(z, G_n)$ be the depths of a point $z \in \mathcal{Z}$ with respect to F_m and G_n respectively, where F_m and G_n are the empirical distribution functions of F and G and $\mathcal{Z} = \mathcal{X} \cup \mathcal{Y}$ is the combined sample of $\mathcal{X}$ and $\mathcal{Y}$. Define

$$DD(F_m, G_n) = \{(D(z, F_m), D(z, G_n)), \forall z \in \mathcal{Z}\}.$$

DD-plot is the plot of points in the set $DD(F_m, G_n)$. To be specific, it is a two dimensional graph of the depth values $D(z, F_m)$ and $D(z, G_n)$. DD-plot is a graphical tool used for comparing two multivariate
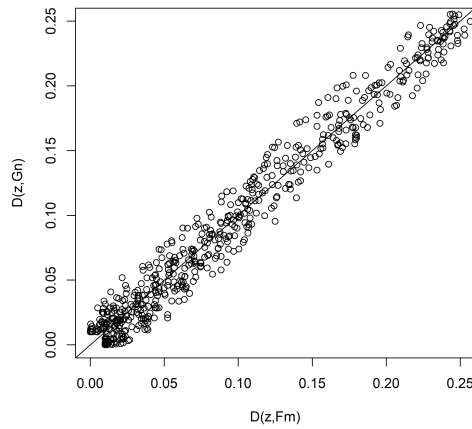
Figure 1: *DD-plots of $F = MN_2((0,0),I_2)$ and $G = MN_2((0,0),I_2)$. DD-plot = depth-versus-depth plot.*

(a)   (b)

Figure 2: *DD-plots of (a) $F = MN_2((0,0),I_2)$ and $G = MN_2((0,0),0.2 * I_2)$ and (b) $F = MN_2((0,0),I_2)$ and $G = MN_2((0,0),0.5 * I_2)$. DD-plot = depth-versus-depth plot.*

samples. The points on the DD-plot are close to the diagonal line if both the samples come from the same distribution (Figure 1). The points then fall below the diagonal line or above the diagonal line if we assume that location parameters are same and the scale parameters of two multivariate distributions are not the same. We observe from Figure 2 and Figure 3 that if the scale parameter corresponding to one of the distributions F or G is large then the depth values with respect to this distribution are also large. In this study, a 'ddalpha' package has been used which is available in R-software (R Core Team, 2018).

3. Data depth based test for multivariate scale difference and F_{product} test

Let $\mathcal{X} = \{X_1, X_2, \ldots, X_m\}$ and $\mathcal{Y} = \{Y_1, Y_2, \ldots, Y_n\}$ be a two data vectors of sizes m and n observed from a continuous distribution with cumulative distribution functions F and G, respectively, where
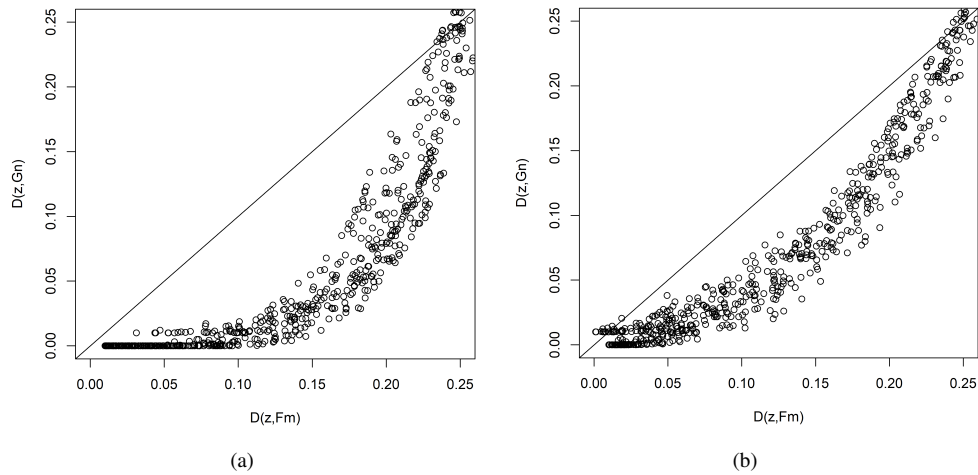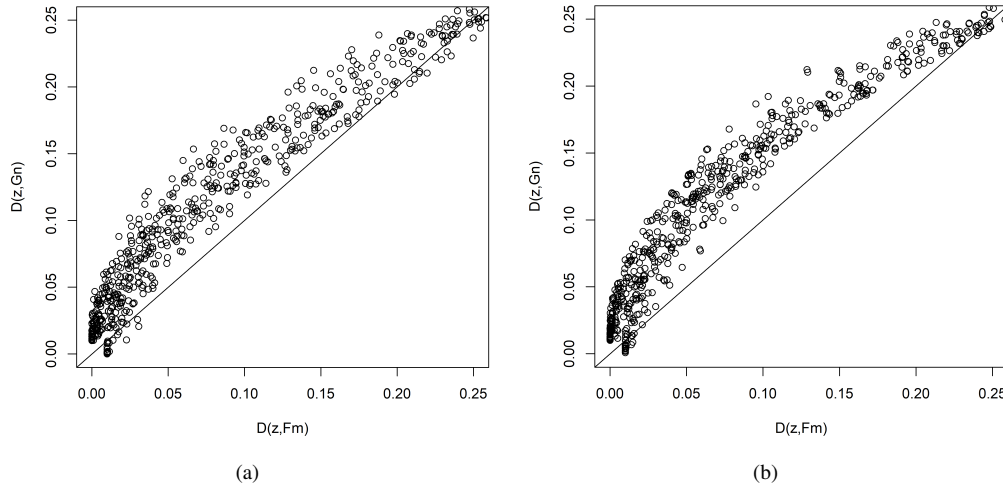
Figure 3: *DD-plots of (a) $F = MN_2((0,0), I_2)$ and $G = MN_2((0,0), 1.5 * I_2)$ and (b) $F = MN_2((0,0), I_2)$ and $G = MN_2((0,0), 2.0 * I_2)$.*

each $X_i \in \mathbb{R}^p$ and $Y_j \in \mathbb{R}^p$, $i = 1, 2, \ldots, m$, $j = 1, 2, \ldots, n$. We assume that F and G are identical except for a possible scale difference. The problem is to test the null hypothesis,

$$H_0 : F(x) = G(x), \quad \forall x = (x_1, x_2, \ldots, x_p)^T,$$

against an alternative hypothesis

$$H_1 : F(x) = G(\sigma x) \quad \forall x = (x_1, x_2, \ldots, x_p)^T,$$

where $\sigma x = (\sigma_1 x_1, \sigma_2 x_2, \ldots, \sigma_p x_p)$ and all elements of $\sigma = (\sigma_1, \sigma_2, \ldots, \sigma_p)$ are positive. It is equivalent to test $H_0 : \sigma = 1$ against $H_1 : \sigma \neq 1$, where 1 is a vector of 1's. Let F_m and G_n denote the empirical distribution functions of F and G, respectively and $D(z, F_m)$ and $D(z, G_n)$ be the depths of a point $z \in \mathcal{Z}$ with respect to F_m and G_n respectively, where $\mathcal{Z} = \{Z_1, Z_2, \ldots, Z_{m+n}\} = \{X_1, X_2, \ldots, X_m, Y_1, Y_2, \ldots, Y_n\}$ is the combined sample of $\mathcal{X}$ and $\mathcal{Y}$.

3.1. Test based on depth based center-outward rank

Liu and Singh (2006) have proposed a nonparametric test to compare the scale parameters of two multivariate distributions using the notion of data depth. This test can be considered as a natural generalization of the Wilcoxon rank-sum test. It is based on the ranks of the depth values of observations $\{Y_1, Y_2, \ldots, Y_n\}$ with respect to the combined sample $\mathcal{Z}$. The test statistic is defined as,

$$R = \sum_{j=1}^{n} \text{rank}(Y_j),$$

where $\text{rank}(Y_j) = \#\{z \in \mathcal{Z} : D(z, \mathcal{Z}) \leq D(Y_j, \mathcal{Z})\}$ and $D(z, \mathcal{Z})$ is the depth of a point z with respect to the combined sample $\mathcal{Z}$. Larger ranks are associated with central observations and smaller ranks are associated with extreme observations. The test rejects the null hypothesis for the smaller value of the test statistic R. The p-value of this test is obtained using the Wilcoxon rank sum table (Hettmansperger, 1984).

3.2. F_{product} test

F_{product} test is used to compare the scales of two multivariate normal distributions. Let $\mathcal{X} = \{X_1, X_2, \ldots, X_m\}$ and $\mathcal{Y} = \{Y_1, Y_2, \ldots, Y_n\}$ be the two independent random samples of size m and n drawn from two multivariate normal populations with means μ_1 and μ_2 and variance-covariance matrices Σ_1 and Σ_2 respectively, where each $X_i \in \mathbb{R}^p$ and $Y_j \in \mathbb{R}^p$, $i = 1, 2, \ldots, m$, $j = 1, 2, \ldots, n$. The notation $|\Sigma|$ indicates the determinant of the variance-covariance matrix Σ. The null hypothesis is to test $H_0 : |\Sigma_1| = |\Sigma_2|$ against $H_1 : |\Sigma_1| < |\Sigma_2|$. Let A_1 and A_2 denote the sample variance-covariance matrices obtained from the samples $\mathcal{X}$ and $\mathcal{Y}$ respectively. The F_{product} test statistic for testing above hypothesis is,

$$F_{\text{prod}} = \frac{(m-1)^{p-1} \prod_{i=1}^{p-1} (n-i-1)|A_2|}{(n-1)^{p-1} \prod_{i=1}^{p-1} (m-i-1)|A_1|}.$$

The test statistic F_{prod} follows the same distribution under H_0 as the distribution of $V_1 V_2 \cdots V_p$, where V_i's are independently distributed as F-distribution with degrees of freedom $(m-i, n-i)$. Test reject the null hypothesis for larger values of the F_{prod}. That is, reject H_0, if $F_{\text{prod}} \geq \hat{F}_\alpha$, where $\hat{F}$ is the distribution of $V_1 V_2 \cdots V_p$ and $\hat{F}_\alpha$ is an upper α^{th} quantile of $\hat{F}$. By using simulation, $\hat{F}$ can be approximated.

4. Proposed two sample tests

It is observed from the DD-plot that when the two distributions F and G are identical, the depth values with respect to F_m and G_n of any point are approximately same. The depth values with respect to F_m and G_n of any point are significantly different if the two distributions F and G are not identical; in addition, the depth of any point with respect to this distribution is also large (small) if one of the distributions F or G has a larger scale (smaller scale). By motivating this fact, we propose two test statistics for testing scale difference based on all $(m+n)^2$ differences between the paired depth values. These test statistics are defined as,

- S_1 test statistic

$$S_1 = \sum_{z \in \mathcal{Z}} \sum_{z^* \in \mathcal{Z}} |D(z, F_m) - D(z^*, G_n)|,$$

- S_2 test statistic

$$S_2 = \sum_{z \in \mathcal{Z}} \sum_{z^* \in \mathcal{Z}} (D(z, F_m) - D(z^*, G_n))^2.$$

Larger scale difference will lead to larger values of S_1 and S_2. Therefore, each test rejects H_0 for the larger values of the test statistic. By using Fisher's permutation approach, the p-values of the proposed tests are obtained. Define

$$P_B^{S_1} = \frac{\sum_{i=1}^B I\left(S_{1i}^* \geq S_{1\text{obs}}\right)}{B},$$

where, B is the total number of permutations of the original combined sample $X \cup Y$, $I(\cdot)$ is an indicator function, $S_{1\text{obs}}$ is the observed value of the S_1 obtained from the observed combined sample and S_{1i}^*

Table 1: Bivariate distributions used for the simulation study

Distribution	Under H_0	Under H_1
Symmetric normal	$N_2(\theta = (0,0), \sigma = (1,1))$	$N_2(\theta = (0,0), \sigma = (c,c))$
Symmetric t	$t_2(\theta = (0,0), \sigma = (1,1), v = 10)$	$t_2(\theta = (0,0), \sigma = (c,c), v = 10)$
Skew normal	$SN_2(\theta = (0,0), \sigma = (1,1), \alpha = (10,4))$	$SN_2(\theta = (0,0), \sigma = (c,c), \alpha = (10,4))$
Skew t	$ST_2(\theta = (0,0), \sigma = (1,1), \alpha = (10,4), v = 10)$	$ST_2(\theta = (0,0), \sigma = (c,c), \alpha = (10,4), v = 10)$
Clayton	$ClaytonGam_2((\beta_1 = 1, \lambda_1 = 1), (\beta_2 = 1, \lambda_2 = 1), \xi = 0.5)$	$ClaytonGam_2((\beta_1 = 1, \lambda_1 = c), (\beta_2 = 1, \lambda_2 = c), \xi = 0.5)$
Frank	$FrankGam_2((\beta_1 = 1, \lambda_1 = 1), (\beta_2 = 1, \lambda_2 = 1), \xi = 0.5)$	$FrankGam_2((\beta_1 = 1, \lambda_1 = c), (\beta_2 = 1, \lambda_2 = c), \xi = 0.5)$
Gumbel	$GumbelGam_2((\beta_1 = 1, \lambda_1 = 1), (\beta_2 = 1, \lambda_2 = 1), \xi = 1.5)$	$GumbelGam_2((\beta_1 = 1, \lambda_1 = c), (\beta_2 = 1, \lambda_2 = c), \xi = 1.5)$

Table 2: Trivariate distributions used for the simulation study

Distribution	Under H_0
Symmetric normal	$N_3(\theta = (0,0,0), \sigma = (1,1,1))$
Symmetric t	$t_3(\theta = (0,0,0), \sigma = (1,1,1), v = 10)$
Skew normal	$SN_3(\theta = (0,0,0), \sigma = (1,1,1), \alpha = (10,4,4))$
Skew t	$ST_3(\theta = (0,0,0), \sigma = (1,1,1), \alpha = (10,4,4), v = 10)$
Clayton	$ClaytonGam_3((\beta_1 = 1, \lambda_1 = 1), (\beta_2 = 1, \lambda_2 = 1), (\beta_3 = 1, \lambda_3 = 1), \xi = 0.5)$
Frank	$FrankGam_3((\beta_1 = 1, \lambda_1 = 1), (\beta_2 = 1, \lambda_2 = 1), (\beta_3 = 1, \lambda_3 = 1), \xi = 0.5)$
Gumbel	$GumbelGam_3((\beta_1 = 1, \lambda_1 = 1), (\beta_2 = 1, \lambda_2 = 1), (\beta_3 = 1, \lambda_3 = 1), \xi = 1.5)$
Distribution	**Under H_1**
Symmetric normal	$N_3(\theta = (0,0,0), \sigma = (c,c,c))$
Symmetric t	$t_3(\theta = (0,0,0), \sigma = (c,c,c), v = 10)$
Skew normal	$SN_3(\theta = (0,0,0), \sigma = (c,c,c), \alpha = (10,4,4))$
Skew t	$ST_3(\theta = (0,0,0), \sigma = (c,c,c), \alpha = (10,4,4), v = 10)$
Clayton	$ClaytonGam_3((\beta_1 = 1, \lambda_1 = c), (\beta_2 = 1, \lambda_2 = c), (\beta_3 = 1, \lambda_3 = c), \xi = 0.5)$
Frank	$FrankGam_3((\beta_1 = 1, \lambda_1 = c), (\beta_2 = 1, \lambda_2 = c), (\beta_3 = 1, \lambda_3 = c), \xi = 0.5)$
Gumbel	$GumbelGam_3((\beta_1 = 1, \lambda_1 = c), (\beta_2 = 1, \lambda_2 = c), (\beta_3 = 1, \lambda_3 = c), \xi = 1.5)$

is the value of the S_1 corresponding to the ith permutation of the combined sample, $i = 1, 2, \ldots, B$. In a similar manner, we can also calculate the p-values $P_B^{S_2}$ as,

$$P_B^{S_2} = \frac{\sum_{i=1}^{B} I\left(S_{2i}^* \geq S_{2\text{obs}}\right)}{B}.$$

5. Simulation study

In this section, the performance of the proposed tests, rank test and parametric F_{product} test, in terms of empirical power, have been evaluated for various symmetric, skewed and copula families of distributions. Table 1 list these distributions. The parameters θ, σ, v, α, and ξ respectively denote the location parameter, scale parameter, degrees of freedom, shape parameter and parameter of the generator of copulas. The notations ClaytonGam, FrankGam, and GumbelGam represent the Clayton, Frank and Gumbel copula sub-families of the distributions with the gamma marginals. We use the Archimedean copula to generate random numbers from distributions. For Clayton and Frank copulas, the parameter of generator ξ is 0.5 and for Gumbel copula, $\xi = 1.5$. For gamma marginals, the shape parameters (β_1, β_2) are equal to $(1, 1)$ and the scale parameters (λ_1, λ_2) are equal to (c, c), where $c = 1.0, 1.2, 1.4, 1.6, 1.8, 2.0$. Details regarding multivariate skewed distributions are provided in Dovoedo and Chakraborti (2015). The results are reported for various values of $\sigma = (c, c)$, where $c = 1.0, 1.2, 1.4, 1.6, 1.8, 2.0$. A distribution F always corresponds to value of $\theta = (0, 0)$ and $\sigma = (1, 1)$ and the distribution G corresponds to $\theta = (0, 0)$ and any value of $\sigma = (c, c)$. We have obtained the performance for sample sizes $m = n = 100$ and $m = n = 150$. Four different depth functions namely SD, MD, HSD, and SPD are used to calculate the depth of observation. The p-values of our proposed

Table 3: Empirical powers of the proposed, rank and F_{product} tests for bivariate normal and bivariate t distributions with sample size $m = n = 100$

Depth	c	Normal				t			
		S_1	S_2	R	F_{product}	S_1	S_2	R	F_{product}
Simplicial	1.0	0.063	0.066	0.059	0.054	0.052	0.053	0.051	0.102
	1.2	0.180	0.170	0.204	0.340	0.156	0.147	0.179	0.379
	1.4	0.506	0.481	0.420	0.770	0.415	0.392	0.369	0.714
	1.6	0.787	0.758	0.631	0.953	0.702	0.674	0.586	0.903
	1.8	0.930	0.915	0.797	0.992	0.872	0.858	0.751	0.977
	2.0	0.986	0.980	0.886	0.999	0.953	0.943	0.858	0.997
Mahalanobis	1.0	0.044	0.046	0.038	0.041	0.050	0.053	0.049	0.096
	1.2	0.235	0.235	0.195	0.359	0.157	0.157	0.158	0.371
	1.4	0.614	0.601	0.404	0.755	0.455	0.454	0.369	0.709
	1.6	0.884	0.877	0.628	0.944	0.726	0.725	0.579	0.903
	1.8	0.981	0.981	0.800	0.995	0.891	0.888	0.747	0.976
	2.0	0.997	0.996	0.898	1.000	0.966	0.967	0.859	0.997
Halfspace	1.0	0.059	0.056	0.056	0.055	0.056	0.057	0.059	0.103
	1.2	0.171	0.155	0.179	0.356	0.153	0.134	0.181	0.376
	1.4	0.452	0.408	0.400	0.742	0.403	0.364	0.379	0.694
	1.6	0.766	0.708	0.630	0.951	0.670	0.636	0.594	0.913
	1.8	0.918	0.881	0.782	0.993	0.862	0.826	0.765	0.979
	2.0	0.979	0.963	0.899	0.999	0.947	0.926	0.861	0.996
Spatial	1.0	0.044	0.043	0.056	0.059	0.046	0.044	0.052	0.091
	1.2	0.196	0.193	0.195	0.363	0.169	0.162	0.174	0.413
	1.4	0.492	0.466	0.411	0.750	0.430	0.409	0.373	0.724
	1.6	0.794	0.765	0.608	0.954	0.697	0.681	0.570	0.901
	1.8	0.935	0.925	0.793	0.993	0.877	0.862	0.734	0.975
	2.0	0.979	0.973	0.891	1.000	0.957	0.947	0.861	0.996

tests are determined using Fisher's permutation principle and we used 500 permutations to calculate p-values. The empirical power of the proposed S_1 and S_2 tests are obtained by the proportion of p-values less than or equal to a specified level of the significance 5%. We used 1,000 simulations to report the power. The performance of the proposed, rank and F_{product} tests is also obtained for trivariate distributions with MD and SPD when the sample size $m = n = 100$. Table 2 provides details regarding trivariate distributions.

Table 3 and Table 4 report the simulation results in terms of the empirical powers for bivariate normal and t distributions with sample sizes $m = n = 100$ and $m = n = 150$. Table 5 and Table 6 show the empirical powers for bivariate skew normal and skew t distributions with sample sizes $m = n = 100$ and $m = n = 150$, respectively. Table 7 and Table 8 give the empirical powers for the copula families of distributions. Table 9 to Table 11 shows the empirical powers for trivariate distributions when the sample size $m = n = 100$. The first row corresponding to depth functions in the Table of empirical power indicates the size of the respective tests. Table 12 provides the ranking of the S_1, S_2, and R tests with respect to empirical power for all depth functions and distributions with the sample size $m = n = 150$. The lower rank in the Table indicates the better power. Table 13 shows the ranking of the depth functions with respect to empirical power for S_1, S_2, and R tests with distributions. Table 12 and Table 13 provide information about the performance of the proposed and rank tests under different depth functions and different distributions. We have observed the same conclusions for the sample size $m = n = 100$ as that for the sample size $m = n = 150$. Therefore, we have reported the conclusions for the bivariate distributions with sample size $m = n = 150$ and trivariate distributions with $m = n = 100$ in the following.

Table 4: Empirical powers of the proposed, rank and F_{product} tests for bivariate normal and bivariate t distributions with sample size $m = n = 150$

Depth	c	Distribution							
		Normal				t			
		S_1	S_2	R	F_{product}	S_1	S_2	R	F_{product}
Simplicial	1.0	0.041	0.046	0.053	0.044	0.050	0.050	0.048	0.106
	1.2	0.269	0.257	0.246	0.504	0.229	0.212	0.240	0.472
	1.4	0.697	0.662	0.551	0.903	0.620	0.583	0.499	0.845
	1.6	0.929	0.909	0.777	0.995	0.869	0.849	0.724	0.961
	1.8	0.994	0.985	0.924	1.000	0.968	0.959	0.870	0.997
	2.0	1.000	0.999	0.973	1.000	0.996	0.995	0.952	1.000
Mahalanobis	1.0	0.038	0.038	0.050	0.046	0.054	0.054	0.040	0.088
	1.2	0.343	0.335	0.246	0.504	0.235	0.233	0.207	0.464
	1.4	0.809	0.804	0.537	0.910	0.606	0.606	0.496	0.843
	1.6	0.979	0.975	0.774	0.992	0.889	0.889	0.745	0.967
	1.8	0.996	0.996	0.903	1.000	0.978	0.978	0.886	0.996
	2.0	1.000	1.000	0.970	1.000	0.993	0.993	0.951	0.999
Halfspace	1.0	0.054	0.056	0.055	0.048	0.051	0.049	0.046	0.096
	1.2	0.255	0.227	0.247	0.505	0.215	0.195	0.216	0.465
	1.4	0.658	0.598	0.544	0.907	0.556	0.514	0.486	0.830
	1.6	0.914	0.875	0.791	0.996	0.847	0.809	0.739	0.973
	1.8	0.984	0.972	0.919	1.000	0.964	0.946	0.896	0.997
	2.0	0.999	0.994	0.973	1.000	0.987	0.984	0.955	1.000
Spatial	1.0	0.055	0.053	0.054	0.049	0.044	0.046	0.040	0.077
	1.2	0.268	0.256	0.253	0.497	0.210	0.199	0.233	0.466
	1.4	0.678	0.651	0.546	0.901	0.583	0.570	0.488	0.845
	1.6	0.914	0.902	0.778	0.992	0.898	0.885	0.742	0.974
	1.8	0.986	0.983	0.921	1.000	0.970	0.966	0.889	0.997
	2.0	0.999	0.998	0.974	1.000	0.992	0.991	0.950	1.000

Table 5: Empirical powers of the proposed, rank and F_{product} tests for bivariate skew normal and bivariate skew t distributions with sample size $m = n = 100$

Depth	c	Distribution							
		Skew normal				Skew t			
		S_1	S_2	R	F_{product}	S_1	S_2	R	F_{product}
Simplicial	1.0	0.055	0.052	0.051	0.059	0.053	0.048	0.052	0.131
	1.2	0.123	0.120	0.154	0.370	0.097	0.096	0.119	0.394
	1.4	0.318	0.313	0.281	0.739	0.237	0.239	0.238	0.689
	1.6	0.547	0.553	0.431	0.928	0.425	0.430	0.353	0.869
	1.8	0.734	0.732	0.569	0.985	0.586	0.600	0.468	0.943
	2.0	0.844	0.858	0.677	0.999	0.704	0.729	0.574	0.983
Mahalanobis	1.0	0.044	0.048	0.050	0.065	0.059	0.055	0.051	0.148
	1.2	0.168	0.166	0.162	0.372	0.129	0.130	0.137	0.408
	1.4	0.428	0.430	0.312	0.751	0.294	0.299	0.267	0.691
	1.6	0.713	0.713	0.485	0.937	0.487	0.494	0.420	0.875
	1.8	0.880	0.883	0.636	0.988	0.678	0.678	0.558	0.953
	2.0	0.948	0.948	0.753	0.998	0.810	0.809	0.673	0.988
Halfspace	1.0	0.048	0.055	0.047	0.063	0.050	0.049	0.053	0.128
	1.2	0.122	0.112	0.147	0.370	0.097	0.094	0.109	0.388
	1.4	0.291	0.267	0.289	0.739	0.214	0.202	0.205	0.689
	1.6	0.493	0.465	0.445	0.928	0.349	0.339	0.335	0.871
	1.8	0.628	0.618	0.578	0.985	0.479	0.480	0.455	0.958
	2.0	0.729	0.713	0.683	0.999	0.572	0.568	0.562	0.991
Spatial	1.0	0.051	0.049	0.054	0.069	0.052	0.051	0.047	0.119
	1.2	0.130	0.132	0.151	0.370	0.117	0.113	0.122	0.388
	1.4	0.375	0.360	0.306	0.739	0.290	0.282	0.236	0.689
	1.6	0.618	0.607	0.471	0.928	0.488	0.482	0.381	0.871
	1.8	0.797	0.786	0.618	0.985	0.672	0.663	0.523	0.958
	2.0	0.903	0.889	0.730	0.999	0.783	0.771	0.641	0.991

Table 6: Empirical powers of the proposed, rank and F_{product} tests for bivariate skew normal and bivariate skew t distributions with sample size $m = n = 150$

Depth	c	Skew normal				Skew t			
		S_1	S_2	R	F_{product}	S_1	S_2	R	F_{product}
Simplicial	1.0	0.055	0.055	0.056	0.062	0.050	0.051	0.052	0.136
	1.2	0.175	0.174	0.183	0.465	0.142	0.137	0.133	0.492
	1.4	0.462	0.464	0.355	0.867	0.346	0.350	0.282	0.800
	1.6	0.724	0.732	0.527	0.985	0.564	0.583	0.458	0.948
	1.8	0.867	0.885	0.710	1.000	0.755	0.781	0.611	0.991
	2.0	0.942	0.947	0.821	1.000	0.822	0.847	0.730	0.999
Mahalanobis	1.0	0.052	0.052	0.052	0.056	0.051	0.051	0.050	0.128
	1.2	0.226	0.228	0.185	0.483	0.144	0.146	0.164	0.486
	1.4	0.622	0.622	0.402	0.867	0.413	0.418	0.345	0.806
	1.6	0.887	0.890	0.622	0.982	0.661	0.669	0.532	0.944
	1.8	0.977	0.979	0.783	0.999	0.858	0.862	0.698	0.989
	2.0	0.996	0.995	0.874	1.000	0.936	0.940	0.808	1.000
Halfspace	1.0	0.061	0.062	0.050	0.066	0.045	0.043	0.049	0.141
	1.2	0.175	0.165	0.175	0.459	0.137	0.126	0.145	0.467
	1.4	0.432	0.400	0.372	0.879	0.294	0.281	0.303	0.801
	1.6	0.656	0.632	0.583	0.986	0.491	0.484	0.467	0.947
	1.8	0.788	0.766	0.708	0.999	0.638	0.634	0.618	0.992
	2.0	0.868	0.849	0.815	1.000	0.717	0.721	0.704	0.998
Spatial	1.0	0.043	0.043	0.054	0.063	0.049	0.053	0.050	0.130
	1.2	0.187	0.184	0.186	0.466	0.153	0.156	0.161	0.472
	1.4	0.533	0.527	0.393	0.872	0.416	0.408	0.318	0.800
	1.6	0.808	0.798	0.603	0.985	0.667	0.662	0.526	0.950
	1.8	0.935	0.922	0.761	1.000	0.834	0.823	0.675	0.987
	2.0	0.979	0.973	0.872	1.000	0.922	0.920	0.798	0.998

Table 7: Empirical powers of the proposed, rank and F_{product} tests for copula families of bivariate distributions with sample size $m = n = 100$

Depth	c	Clayton gamma				Frank gamma				Gumbel gamma			
		S_1	S_2	R	F_{product}	S_1	S_2	R	F_{product}	S_1	S_2	R	F_{product}
Simplicial	1.0	0.056	0.058	0.061	0.215	0.056	0.055	0.051	0.193	0.045	0.044	0.050	0.208
	1.2	0.100	0.112	0.115	0.676	0.106	0.108	0.086	0.669	0.103	0.113	0.112	0.670
	1.4	0.234	0.266	0.207	0.938	0.209	0.264	0.140	0.947	0.248	0.284	0.206	0.934
	1.6	0.382	0.460	0.297	0.994	0.313	0.425	0.218	0.995	0.395	0.475	0.299	0.995
	1.8	0.488	0.591	0.385	1.000	0.365	0.529	0.267	0.999	0.506	0.622	0.405	1.000
	2.0	0.533	0.664	0.457	1.000	0.364	0.582	0.323	1.000	0.555	0.694	0.492	1.000
Mahalanobis	1.0	0.046	0.047	0.048	0.207	0.052	0.056	0.058	0.188	0.059	0.059	0.046	0.211
	1.2	0.116	0.130	0.125	0.679	0.133	0.134	0.118	0.669	0.132	0.144	0.158	0.668
	1.4	0.378	0.384	0.255	0.933	0.378	0.378	0.210	0.928	0.398	0.425	0.317	0.939
	1.6	0.683	0.674	0.388	0.992	0.665	0.655	0.315	0.994	0.692	0.715	0.476	0.993
	1.8	0.871	0.849	0.509	0.999	0.853	0.829	0.429	0.999	0.878	0.890	0.609	1.000
	2.0	0.947	0.933	0.603	1.000	0.942	0.917	0.524	1.000	0.955	0.958	0.717	1.000
Halfspace	1.0	0.054	0.053	0.047	0.211	0.052	0.050	0.050	0.198	0.046	0.052	0.046	0.185
	1.2	0.090	0.101	0.111	0.676	0.091	0.099	0.085	0.669	0.093	0.105	0.114	0.670
	1.4	0.175	0.207	0.190	0.938	0.158	0.196	0.151	0.947	0.198	0.238	0.204	0.934
	1.6	0.240	0.291	0.284	0.994	0.196	0.282	0.213	0.995	0.302	0.348	0.304	0.995
	1.8	0.263	0.337	0.369	1.000	0.207	0.302	0.268	0.999	0.341	0.416	0.403	1.000
	2.0	0.275	0.343	0.436	1.000	0.211	0.306	0.335	1.000	0.351	0.451	0.498	1.000
Spatial	1.0	0.054	0.053	0.056	0.196	0.053	0.053	0.056	0.196	0.061	0.059	0.059	0.202
	1.2	0.160	0.161	0.141	0.676	0.176	0.178	0.129	0.660	0.158	0.160	0.151	0.670
	1.4	0.411	0.411	0.278	0.938	0.409	0.421	0.258	0.938	0.422	0.426	0.312	0.934
	1.6	0.645	0.646	0.431	0.994	0.650	0.649	0.367	0.995	0.678	0.681	0.477	0.995
	1.8	0.806	0.806	0.575	1.000	0.810	0.801	0.506	0.999	0.839	0.838	0.624	1.000
	2.0	0.894	0.881	0.694	1.000	0.881	0.870	0.604	1.000	0.916	0.913	0.748	1.000

Table 8: Empirical powers of the proposed, rank and F_{product} tests for copula families of bivariate distributions with sample size $m = n = 150$

Depth	c	Distribution											
		Clayton gamma				Frank gamma				Gumbel gamma			
		S_1	S_2	R	F_{product}	S_1	S_2	R	F_{product}	S_1	S_2	R	F_{product}
Simplicial	1.0	0.057	0.055	0.052	0.200	0.049	0.048	0.052	0.214	0.052	0.052	0.056	0.193
	1.2	0.131	0.144	0.137	0.752	0.108	0.131	0.121	0.775	0.137	0.154	0.137	0.774
	1.4	0.318	0.395	0.251	0.988	0.254	0.342	0.206	0.987	0.309	0.378	0.267	0.983
	1.6	0.493	0.586	0.411	1.000	0.326	0.476	0.308	1.000	0.478	0.582	0.418	0.998
	1.8	0.550	0.678	0.504	1.000	0.341	0.552	0.394	1.000	0.542	0.680	0.532	1.000
	2.0	0.554	0.704	0.605	1.000	0.277	0.536	0.451	1.000	0.579	0.751	0.637	1.000
Mahalanobis	1.0	0.049	0.051	0.048	0.212	0.061	0.060	0.054	0.203	0.054	0.053	0.056	0.212
	1.2	0.199	0.209	0.177	0.792	0.187	0.203	0.159	0.765	0.181	0.214	0.198	0.774
	1.4	0.573	0.573	0.335	0.981	0.567	0.579	0.296	0.981	0.542	0.591	0.381	0.977
	1.6	0.878	0.866	0.483	0.999	0.890	0.865	0.438	1.000	0.860	0.874	0.602	0.999
	1.8	0.966	0.962	0.644	1.000	0.976	0.965	0.587	1.000	0.970	0.972	0.761	1.000
	2.0	0.997	0.991	0.752	1.000	0.994	0.989	0.684	1.000	0.998	0.996	0.857	1.000
Halfspace	1.0	0.047	0.048	0.052	0.194	0.049	0.048	0.059	0.208	0.043	0.047	0.058	0.208
	1.2	0.127	0.137	0.140	0.787	0.103	0.123	0.112	0.771	0.140	0.147	0.135	0.752
	1.4	0.231	0.259	0.247	0.979	0.186	0.253	0.198	0.976	0.273	0.324	0.254	0.982
	1.6	0.312	0.388	0.376	1.000	0.226	0.341	0.290	0.999	0.395	0.456	0.406	1.000
	1.8	0.345	0.428	0.495	1.000	0.199	0.353	0.374	1.000	0.404	0.508	0.536	1.000
	2.0	0.356	0.431	0.586	1.000	0.206	0.358	0.434	1.000	0.412	0.548	0.636	1.000
Spatial	1.0	0.059	0.063	0.051	0.215	0.057	0.055	0.052	0.215	0.053	0.053	0.054	0.206
	1.2	0.218	0.226	0.175	0.764	0.218	0.226	0.160	0.772	0.220	0.225	0.185	0.763
	1.4	0.574	0.572	0.340	0.986	0.573	0.588	0.314	0.983	0.588	0.598	0.388	0.978
	1.6	0.815	0.818	0.557	1.000	0.801	0.797	0.489	1.000	0.846	0.842	0.613	0.998
	1.8	0.936	0.928	0.726	1.000	0.903	0.899	0.654	1.000	0.953	0.949	0.787	1.000
	2.0	0.968	0.963	0.830	1.000	0.965	0.963	0.780	1.000	0.982	0.979	0.877	1.000

Power performance for bivariate distributions with sample size $m = n = 150$

The proposed S_1 and S_2 tests perform better than the rank test for all the symmetric as well as the skewed families of distributions. In the case of a smaller scale change up to 1.2, the rank test works better than the proposed tests for some of the distributions and depth functions. F_{product} test is more powerful than other remaining tests for symmetric normal and skew normal distributions with all depth functions. However, the sizes of the F_{product} test are large for all other distributions and depth functions.

The power performance of the proposed tests and rank test for a normal distribution is better than all other distributions. The proposed S_1 and S_2 tests have greater power for Mahalanobis depth while for a halfspace depth, proposed S_1 and S_2 tests have less power for all the distributions. In the case of skewed and copula families of distributions, the proposed and rank tests have greater powers for Mahalanobis and spatial depths than the remaining depth functions. The proposed and rank tests have larger empirical powers for skew normal distributions than skew t distributions. The depth based tests also perform well for Gumbel distribution with most of the depth functions among all these copula families of distributions.

Power performance for trivariate distributions with sample size $m = n = 100$

Table 9 shows that the F_{product} test is more powerful than remaining tests for normal distribution; however, the proposed S_1 and S_2 tests have better power than the rank test in the case of t distribution. The proposed and rank tests work better for Mahalanobis depth than the spatial depth. Table 10 shows that the S_1 and S_2 perform well for skew normal and skew t distributions. The rank test

Table 9: Empirical powers of the proposed, rank and F_{product} tests for trivariate normal and trivariate t distributions with sample size $m = n = 100$

Depth	c	Distribution							
		Normal				t			
		S_1	S_2	R	F_{product}	S_1	S_2	R	F_{product}
Mahalanobis	1.0	0.052	0.048	0.046	0.053	0.050	0.049	0.056	0.151
	1.2	0.340	0.337	0.258	0.585	0.232	0.236	0.220	0.569
	1.4	0.788	0.782	0.565	0.940	0.586	0.588	0.495	0.880
	1.6	0.971	0.964	0.811	0.995	0.847	0.847	0.743	0.976
	1.8	0.997	0.997	0.933	1.000	0.957	0.956	0.886	0.996
	2.0	1.000	1.000	0.976	1.000	0.989	0.989	0.952	1.000
Spatial	1.0	0.049	0.051	0.044	0.051	0.047	0.048	0.049	0.152
	1.2	0.301	0.285	0.263	0.585	0.226	0.218	0.221	0.560
	1.4	0.725	0.711	0.567	0.940	0.583	0.574	0.484	0.878
	1.6	0.943	0.930	0.807	0.995	0.855	0.846	0.719	0.977
	1.8	0.993	0.992	0.934	1.000	0.972	0.968	0.880	0.997
	2.0	0.999	0.999	0.975	1.000	0.995	0.993	0.949	1.000

Table 10: Empirical powers of the proposed, rank and F_{product} tests for trivariate skew normal and trivariate skew t distributions with sample size $m = n = 100$

Depth	c	Distribution							
		Skew normal				Skew t			
		S_1	S_2	R	F_{product}	S_1	S_2	R	F_{product}
Mahalanobis	1.0	0.052	0.050	0.049	0.113	0.050	0.050	0.054	0.197
	1.2	0.251	0.237	0.239	0.580	0.155	0.153	0.196	0.544
	1.4	0.624	0.618	0.507	0.920	0.393	0.387	0.430	0.858
	1.6	0.877	0.856	0.746	0.998	0.664	0.649	0.659	0.962
	1.8	0.968	0.956	0.891	1.000	0.803	0.794	0.801	0.995
	2.0	0.991	0.984	0.957	1.000	0.899	0.889	0.909	1.000
Spatial	1.0	0.047	0.043	0.051	0.097	0.056	0.057	0.051	0.178
	1.2	0.239	0.230	0.220	0.588	0.178	0.179	0.185	0.561
	1.4	0.621	0.603	0.499	0.929	0.474	0.466	0.422	0.844
	1.6	0.890	0.878	0.747	0.996	0.726	0.714	0.626	0.958
	1.8	0.977	0.972	0.890	1.000	0.899	0.889	0.804	0.995
	2.0	0.998	0.996	0.960	1.000	0.969	0.963	0.899	1.000

gives comparable power to the proposed tests for the Mahalanobis depth. Power performance for the Mahalanobis and spatial depth works equally well. By observing Table 9 and Table 10, proposed and rank tests have better power for normal distribution than the remaining distributions with Mahalanobis depth and spatial depth.

Table 11 shows that the rank test performs well for copula families of distribution with Mahalanobis depth; however, the proposed tests are more powerful than the rank test in the case of spatial depth. The proposed tests have better power for spatial depth than the Mahalanobis depth. Also, the power performance of the proposed and rank tests also works well for the Gumbel copula than the Clayton and Frank copula.

It is seen that the proposed S_1 and S_2 tests perform relatively well than the rank test. S_1 and S_2 tests consider all the pairwise differences between the depth values of two multivariate distributions since the difference in scale parameters indicate a difference in the depth values. Therefore, S_1 and S_2 tests capture the variation between the depth values. The empirical power of the discussed tests increases and the size of these tests is close to the significance level of 5% as the sample size increases from $m = n = 100$ to $m = n = 150$.

Table 11: Empirical powers of the proposed, rank and F_{product} tests for copula families of trivariate distributions with sample size $m = n = 100$

Depth	c	Clayton gamma				Frank gamma				Gumbel gamma			
		S_1	S_2	R	F_{product}	S_1	S_2	R	F_{product}	S_1	S_2	R	F_{product}
Mahalanobis	1.0	0.060	0.060	0.048	0.249	0.060	0.060	0.051	0.246	0.055	0.055	0.054	0.253
	1.2	0.122	0.123	0.204	0.806	0.124	0.128	0.206	0.826	0.173	0.173	0.284	0.802
	1.4	0.319	0.317	0.483	0.986	0.288	0.277	0.458	0.991	0.461	0.446	0.617	0.991
	1.6	0.518	0.467	0.694	1.000	0.464	0.392	0.680	0.999	0.730	0.685	0.831	0.999
	1.8	0.652	0.554	0.850	1.000	0.598	0.474	0.833	1.000	0.881	0.830	0.952	1.000
	2.0	0.737	0.593	0.931	1.000	0.656	0.502	0.924	1.000	0.945	0.889	0.989	1.000
Spatial	1.0	0.058	0.061	0.053	0.257	0.048	0.049	0.054	0.233	0.052	0.049	0.054	0.255
	1.2	0.237	0.233	0.216	0.828	0.220	0.215	0.186	0.819	0.267	0.262	0.271	0.806
	1.4	0.628	0.613	0.499	0.990	0.583	0.569	0.430	0.987	0.658	0.648	0.585	0.988
	1.6	0.880	0.857	0.715	1.000	0.833	0.810	0.660	1.000	0.910	0.897	0.828	1.000
	1.8	0.962	0.949	0.872	1.000	0.950	0.925	0.854	1.000	0.979	0.975	0.930	1.000
	2.0	0.992	0.985	0.939	1.000	0.981	0.963	0.923	1.000	0.997	0.995	0.985	1.000

Table 12: Ranking of the S_1, S_2, and R tests for all depths and distributions with sample size $m = n = 150$

Depth	Normal			t			Skew normal			Skew t			Clayton gamma			Frank gamma			Gumbel gamma		
	S_1	S_2	R	S_1	S_2	R	S_1	S_2	R	S_1	S_2	R	S_1	S_2	R	S_1	S_2	R	S_1	S_2	R
SD	1	2	3	1	2	3	2	1	3	2	1	3	2	1	3	2	1	3	2	1	3
MD	1	2	3	1	2	3	2	1	3	2	1	3	1	2	3	1	2	3	2	1	3
HSD	1	2	3	1	2	3	1	2	3	1	2	3	3	2	1	3	1	2	2	1	3
SPD	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3

SD = simplicial depth; MD = Mahalanobis depth; HSD = Halfspace depth; SPD = Spatial depth.

Table 13: Ranking of the SD, MD, HSD, and SPD for all tests and distributions with sample size $m = n = 150$

Depth	Normal			t			Skew normal			Skew t			Clayton gamma			Frank gamma			Gumbel gamma		
	S_1	S_2	R	S_1	S_2	R	S_1	S_2	R	S_1	S_2	R	S_1	S_2	R	S_1	S_2	R	S_1	S_2	R
SD	2	2	2	2	2	1	3	3	4	3	3	4	3	3	3	3	3	3	3	3	3
MD	1	1	1	1	1	4	1	1	1	1	1	1	1	1	2	1	1	2	1	1	2
HSD	4	4	3	4	4	1	4	4	3	4	4	3	4	4	4	4	4	4	4	4	4
SPD	3	3	4	3	3	3	2	2	2	1	2	2	2	2	1	2	1	1	1	1	1

SD = simplicial depth; MD = Mahalanobis depth; HSD = Halfspace depth; SPD = Spatial depth.

6. Application to real-life data

In this section, the applicability of the proposed tests is verified through real-life data. We analyzed the dataset used by Jolicoeur and Mosimann (1960) for further evaluation of the proposed tests. Jolicoeur and Mosimann (1960) studied this dataset for evaluating the applicability of the principal component analysis to size and shape variation in groups of female and male turtles. This dataset contains three measurements in millimeter on the three characteristics namely carapace length, carapace width and carapace height of 48 turtles. Out of 48 turtles, 24 are female turtles and the remaining 24 are male turtles. These are two populations (F and G) corresponding to female and male turtles respectively.

We are interested in testing the scale difference between these two populations. For female and male turtle populations, the multivariate Shapiro test for normality gives p-value 0.1626 and 0.01551 respectively. Therefore, tests based on the assumption of multivariate normality are not appropriate in this case. We construct the DD-plot (Figure 4(a)) for this data using Mahalanobis depth function.
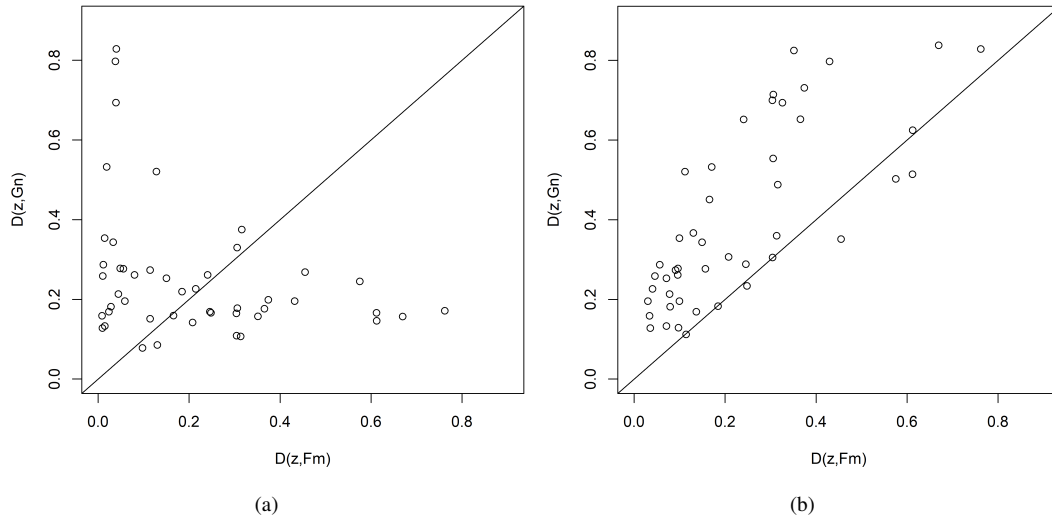
Figure 4: *DD-plot for the real life data (a) before centering data and (b) after centering data.*

Table 14: *p*-values of the proposed and rank tests

Test	S_1	S_2	R
p-value	0.00125	0.00111	0.00172

We first center the data to have the same location at their respective deepest points since the DD-plot indicates location differences between the female and male turtle populations. This can be done by subtracting the deepest point from each data point in the dataset. After centering, the DD-plot (Figure 4(b)) for the same shows the male turtle population has a larger scale than the female turtle population. We also use proposed tests to evaluate if there is a difference in the scale parameters of the distribution of female and male turtles. Table 14 provides the *p*-values for the proposed and rank tests based on $B = 100000$ permutations as well as shows that all the *p*-values are very small. This indicates that female and male turtle populations do not have the same scale parameters.

7. Conclusion

A notion of data depth is used for testing scale parameters of the two multivariate distributions. The proposed tests are nonparametric in nature. The tests are based on all possible pairwise differences between the depth values. The power performance of the proposed tests is conducted for a few symmetric and skewed multivariate distributions along with four depth functions. Power performance is also reported for the large sample size 150 and a higher dimension. The simulation study reveals that the proposed tests work better than the existing depth based test for symmetric as well as skewed multivariate distributions; however, the parametric F_{product} test does not cope well for some non-normal distributions. In addition, the proposed and rank tests perform well for copula families of distributions. A notion of data depth is most commonly used in recent years for nonparametric multivariate analysis. A center-outward ranking produced by data depth is used in various statistical methods such as nonparametric multivariate testing procedures, designing multivariate control charts, and building a multivariate regression model.

Acknowledgments

Bibby JM, Kent J, and Mardia K (1979). *Multivariate Analysis*, Academic Press, London

Box GEP (1949). A general distribution theory for a class of likelihood criteria, *Biometrika*, **36**, 317–346.

Chavan AR and Shirke DT (2016). Nonparametric tests for testing equality of location parameters of two multivariate distributions, *Electronic Journal of Applied Statistical Analysis*, **9**, 417–432.

Chavan AR and Shirke DT (2019). Simultaneously testing for location and scale parameters of two multivariate distributions, *Revista Colombiana de Estadística*, **42**, 185–208.

Chenouri S (2004). *Multivariate robust nonparametric inference based on data depth* (PhD thesis), University of Waterloo, Waterloo.

Chenouri S and Small CG (2012). A nonparametric multivariate multisample test based on data depth, *Electronic Journal of Statistics*, **6**, 760–782.

Dovoedo YH and Chakraborti S (2015). Power of depth-based nonparametric tests for multivariate locations, *Journal of Statistical Computation and Simulation*, **85**, 1987–2006.

Fraiman R, Meloche J, García-Escudero LA, *et al.* (1999). Multivariate L-estimation, *Test*, **8**, 255–317.

Hettmansperger T (1984). *Statistical Inference Based on Ranks*, Wiley, New York.

Jolicoeur P and Mosimann JE (1960). Size and shape variation in the painted turtle. A principal component analysis, *Growth*, **24**, 339–354.

Li J, Ban J, and Santiago LS (2011). Nonparametric tests for homogeneity of species assemblages: a data depth approach, *Biometrics*, **67**, 1481–1488.

Li J, and Liu RY (2004). New nonparametric tests of multivariate locations and scales using data depth, *Statistical Science*, **19**, 686–696.

Li J and Liu RY (2016). New Nonparametric tests for comparing multivariate scales using data depth. In *Robust Rank-Based and Nonparametric Methods*, 209–226, Springer.

Liu RY (1990). On a notion of data depth based on random simplices, *The Annals of Statistics*, **18**, 405–414.

Liu RY, Parelius JM, and Singh K (1999). Multivariate analysis by data depth: descriptive statistics, graphics and inference, (with discussion and a rejoinder by Liu and Singh), *The Annals of Statistics*, **27**, 783–858.

Liu RY and Singh K (1993). A quality index based on data depth and multivariate rank tests, *Journal of the American Statistical Association*, **88**, 252–260.

Liu RY and Singh K (2006). Rank tests for multivariate scale difference based on data depth, *DIMACS Series in Discrete Mathematics and Theoretical Computer Science*, **72**, 17–35.

Mahalanobis PC (1936). On the generalized distance in statistics. In *Proceedings of the National Institute of Sciences (India)*, **2**, 49–55.

Oja H (1983). Descriptive statistics for multivariate distributions, *Statistics & Probability Letters*, **1**, 327–332.

Pawar SD and Shirke DT (2019). Nonparametric tests for multivariate locations based on data depth, *Communications in Statistics-Simulation and Computation*, **48**, 753–776.

R Core Team (2018). *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, https://www.R-project.org/.

Rousson V (2002). On distribution-free tests for the multivariate two-sample location-scale model, *Journal of Multivariate Analysis*, **80**, 43–57.

Serfling R (2002). A depth function and a scale curve based on spatial quantiles, *Statistical Data Analysis Based on the L1-Norm and Related Methods*, 25–38, Springer.

Shirke, DT and Khorate, SD (2017). Power comparison of data depth-based nonparametric tests for testing equality of locations, *Journal of Statistical Computation and Simulation*, **87**, **8**, 1489–1497.

Singh K (1991). *A notion of majority depth* (Technical report), Rutgers University.

Tukey JW (1975). Mathematics and the picturing of data. In *Proceedings of the International Congress of Mathematicians, Vancouver, 1975*, **2**, 523–531.

Zuo Y and Serfling R (2000). General notions of statistical depth function, *Annals of Statistics*, **28**, 461–482.