Scientific
Research
Publishing

A Simulation Based Comparison of Correlation Coefficients with Regard to Type I Error Rate and Power

Elif Tuğran[1], Mehmet Kocak[2], Hamit Mirtagioğlu[3], Soner Yiğit[1], Mehmet Mendes[1*]

[1]Canakkale Onsekiz Mart University, Canakkale, Turkey
[2]University of Tennessee Health Science Center, Memphis, USA
[3]Yüzüncüyıl University, Van, Turkey
Email: [*]mmendes@comu.edu.tr

Abstract

In this simulation study, five correlation coefficients, namely, Pearson, Spearman, Kendal Tau, Permutation-based, and Winsorized were compared in terms of Type I error rate and power under different scenarios where the underlying distributions of the variables of interest, sample sizes and correlation patterns were varied. Simulation results showed that the Type I error rate and power of Pearson correlation coefficient were negatively affected by the distribution shapes especially for small sample sizes, which was much more pronounced for Spearman Rank and Kendal Tau correlation coefficients especially when sample sizes were small. In general, Permutation-based and Winsorized correlation coefficients are more robust to distribution shapes and correlation patterns, regardless of sample size. In conclusion, when assumptions of Pearson correlation coefficient are not satisfied, Permutation-based and Winsorized correlation coefficients seem to be better alternatives.

Keywords

Correlation Coefficient, Robustness, Type I Error Rate, Test Power, Non-Normality, Simulation

1. Introduction

In practice, researchers are often interested in investigating the linear association between two variables, where Pearson-moment correlation coefficient (r) is probably the most commonly used method [1]-[4]. This correlation

---

[*]Corresponding author.

coefficient, however, is only valid under the assumptions of bivariate normality, and linear relationship between and homoscedasticity of the two variables [5]-[8]. It is also highly sensitive to outliers as reported by Wilcox [5] that a single outlier can mask strong relationship between two variables. In practice, these assumptions may not be satisfied, which results in a decrease in power and a deviation of the type I error rate from the nominal level [9]-[12]. The issue of non-normality, a commonly recommended approach is to normalize the non-normal variables using an appropriate transformation and then compute Pearson correlation coefficient on the transformed data. However, in practice, transformation approach does not give satisfactory results in many experimental conditions [13] [14]. Various alternatives have been proposed, such as Spearman-Rank, Kendall-Tau, non-normal transformation based Pearson correlation coefficient, Winsorized and Permutation-based correlation coefficients when assumptions of Pearson-Moment correlation are not satisfied [5] [15] [16], among which Spearman-Rank correlation coefficient is the next most commonly utilized approach in practice while the Winsorized and Permutation-based correlation coefficients may be even more appropriate as they both are not highly affected by distribution shape and outliers [2] [12] [16]-[19]. However, the relative performance of these different methods needs to be explored further.

In this study, we compare the performance of the Pearson-moment correlation coefficient with that of the alternatives, namely, Spearman-Rank, Kendall-Tau, Winsorized and the Permutation-based correlation coefficient under different experimental conditions through extensive simulations. In Section 2, we provide a brief description of the correlation coefficients we are comparing. In Section 3, we describe those experimentation condition, followed by the results of these simulations in Section 4. We end with a practical real-life example and discussions.

2. Materials and Methods

2.1. Simulation Setting

Three different bivariate distributions, namely, bivariate normal distribution, bivariate *t*-distribution with 10 degree of freedom, and bivariate log-normal distribution (skewness = 2.1), were considered, where the choices of bivariate *t* and log-normal distributions represents different degrees of violations of bivariate normality. All three distributions were generated as suitable transformations from a standard bivariate normal distribution with a given covariance structure.

The bivariate normal distributions $\left(X^{t}, Y^{t}\right) \sim BVN\left(\mu_{x}, \mu_{y}, \sigma_{x}^{2}, \sigma_{y}^{2}, \rho_{xy}\right)$ were generated according to the following transformation:

$$\begin{pmatrix} X^{t} \\ Y^{t} \end{pmatrix} = \begin{pmatrix} \mu_{x} \\ \mu_{y} \end{pmatrix} + \left[\Sigma \begin{pmatrix} X \\ Y \end{pmatrix} \right]^{0.5}$$

where $\begin{pmatrix} X \\ Y \end{pmatrix} \sim N\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right]$ and $\Sigma = \begin{pmatrix} \sigma_{x}^{2} & \rho\sigma_{x}\sigma_{y} \\ \rho\sigma_{x}\sigma_{y} & \sigma_{y}^{2} \end{pmatrix}$.

The bivariate *t*-distribution with 10 d.f. $\left(X_{T}, Y_{T}\right) \sim t\left(\rho_{xy}, df\right)$ was generated as follows:

$$\begin{pmatrix} X_{T} \\ X_{T} \end{pmatrix} = \begin{pmatrix} X^{t} \\ Y^{t} \end{pmatrix} \left[\left(\frac{df}{\left(\frac{u}{v} \right)} \right) \right]^{0.5}$$

where, *u* and *v* are independent variables generated from the gamma distribution.

The log-normal distribution was generated as an exponential transformation of the bivariate normal distribution as follows:

$$e^{\left(X^{t}, Y^{t}\right)} = \left(X_{\log-n}, Y_{\log-n}\right) \sim LogNormal\left(\mu_{x}, \mu_{y}, \sigma_{x}^{2}, \sigma_{y}^{2}, \rho_{xy}\right) \text{ [20]-[23]}.$$

For each experimentation condition with sample sizes of 5, 10, 20, 30, 50, and 100, and the true correlation coefficients of 0, 0.2, 0.4, 0.6 and 0.8, 50,000 random samples were generated. Naturally, the experimental conditions with the true Pearson's correlation coefficient of 0 (zero) correspond to the absence of any relations be-

tween two variables; thus, the rejection of the null hypothesis that $\rho = 0$ corresponds to the type I error rate while the rejection rate is the test power when $\rho \neq 0$.

The type I error rate and test power are estimated as the rejection proportion based on the nominal alpha level of 0.05; that is, the rejection rate (R) is calculated as

$$R = \frac{I_{\{P^* < \alpha\}}}{N}$$

where I is the indicator function, P^*, is the permutation p-value, and N is the number of simulations. IMSL library of Fortran Power Station Developer was used for all simulations.

2.2. Correlation Coefficients

2.2.1. Pearson-Moment Correlation Coefficient

Given a bivariate data set of size n, $(x_1, y_1), (x_2, y_2), \cdots, (x_n, y_n)$, the sample Pearson-moment correlation coefficient, r_{xy}, between variables X and Y is defined by the formula

$$r_{xy} = \frac{\sum d_x d_y}{\sqrt{\left(\sum d_x^2\right)\left(\sum d_y^2\right)}}$$

where $\sum d_x d_y$ is the sum of products of the departures from the means, $\sum d_x^2$ and $\sum d_y^2$ are the sum of squares of X and Y variables, respectively [24].

2.2.2. Spearman-Rank Correlation Coefficient

Spearman's rank correlation provides a distribution free test of independence between two variables. Let $R(x)$ and $R(y)$ denote the ranks of a pair of variables (X and Y). In this case, Spearman-Rank correlation coefficient is calculated as

$$r_s = \frac{\sum_{i=1}^{n} R(x_i) R(y_i) - n\left(\frac{n+1}{2}\right)^2}{\sqrt{\left(\sum_{i=1}^{n} R(x_i)^2 - n\left(\frac{n+1}{2}\right)^2\right)\left(\sum_{i=1}^{n} R(y_i)^2 - n\left(\frac{n+1}{2}\right)^2\right)}} \quad [24].$$

2.2.3. Winsorized Correlation

It is well known that the Pearson correlation coefficient is sensitive to outliers. Wilcox (2001) proposed that one of the ways of handling outliers is to compute the Winsorized correlation coefficient, which is computed after the k smallest observations are replaced by the $(k + 1)^{st}$ smallest observation, and the k largest observations are replaced by the $(k + 1)^{st}$ largest observation. Therefore, the observations are winsorizedat each end of both X and Y [5] [18].

Let $(X_1, Y_1), (X_2, Y_2), \cdots, (X_n, Y_n)$ be a random sample from any bivariate distribution. We winsorize X_i to W_i and Y_i to V_i as described above. Then, the Winsorized correlation coefficient r_w, is computed same as the Pearson moment correlation using the Winsorized data as follows:

$$r_w = \frac{\sum \left(W_i - \bar{X}_W\right)\left(V_i - \bar{Y}_W\right)}{\sqrt{\sum \left(W_i - \bar{X}_W\right)^2 \sum \left(V_i - \bar{Y}_W\right)^2}} = \frac{\sum d_W d_V}{\sqrt{\sum d_W^2 \sum d_V^2}}$$

where $\bar{X}_W$ and $\bar{Y}_W$ are the Winsorized means of X and Y variables, respectively Let γ denote the Winsorized percent and define $g = \gamma n$; then, r_w is distributed as t-distribution with (n-2g-2) d.f. [18].

2.2.4. Permutation-Based Pearson Correlation

Principals of permutation tests consist of the following four basic steps:

1) Calculate the Pearson correlation coefficient, r_{xy}, using the original data set.
2) Find all possible permutations as: n!; if n is too large, take a large sample, say, 10,000, from all possible

permutations.

3) Re-calculate the Pearson correlation coefficient for each permutation sample (r_{pxy}), and

4) Compare the absolute value of the correlation coefficients from the permuted samples with the absolute value of the correlation coefficient from the original sample. The permutation-based *p*-value is computed by counting the number of times that correlation coefficients (r_{pxy}) calculated from all permuted samples are greater than or equal to calculated Pearson correlation coefficient (r_{xy}) and dividing it by the total number of permutations used, which is shown as follows:

$$P \cong \frac{\text{Number of simulation with } |r_{pxy}| \geq |r_{xy}|}{\text{Total number of simulation}} \quad \text{[25] [26]}.$$

In our simulations, we used 10,000 permutations to obtain the permutation *p*-values to evaluate the Type-1 error and power.

2.2.5. Kendall-Tau Correlation

Let $(X_1, Y_1), (X_2, Y_2), \cdots, (X_n, Y_n)$ be a set of joint observations from two random variables, *X* and *Y*. Any pair of observations (X_i, Y_i) and (X_j, Y_j) are called to be concordant if the ranks for both pairs agree (*i.e.* if both $X_i > X_j$ and $Y_i > Y_j$ or if both $X_i < X_j$ and $Y_i < Y_j$), otherwise discordant if $X_i > X_j$ and $Y_i < Y_j$ or if $X_i < X_j$ and $Y_i > Y_j$.

Let *C* and *D* denote the number concordant and discordant pairs, respectively. The Kendall Tau coefficient is defined as:

$$r_{tau} = (C - D) / (n(n-1)/2) \quad \text{[27]}.$$

3. Results

The empirical type I error rates and power for Pearson-moment, Spearman-Rank, Kendal-Tau, Permutation-based and Winsorized correlation coefficients are presented in **Tables 1-3**. When samples are taken from bivariate normal distribution, which is the ideal case for Pearson-moment correlation coefficient, except for very small sample sizes such as *n* = 5, all correlation coefficients retain the type I error rate at the nominal. When the type I error rate and power were evaluated together Pearson, Winsorized and Permutation-based correlation coefficients are more robust when compared to the Spearman-Rank and Kendall-Tau correlation coefficients.

Empirical the type I error rates for bivariate *t*-distribution with 10 d.f. are given in **Table 2**. As seen in **Table 2**, the type I error rates are very similar to obtained under bivariate normality presented in **Table 1**, which is not surprising because bivariate *t*-distribution with 10 degrees of freedom is quite close to the bivariate normal distribution. Permutation-based and Winsorized correlation coefficients are the best especially when sample sizes are small (n ≤ 10) under these experimental conditions as they retain the type-I error rate while have higher power than the other approaches. On the other hand, as sample size increases, all correlations coefficient tend to give similar results as expected.

The empirical type I error rate and power of the correlation coefficients under bivariate lognormal distribution are presented in **Table 3**, where it is obvious that the Pearson correlation coefficient is highly negatively affected by the distribution shape. The empirical the type I error rate for Pearson correlation coefficient is estimated to be greater than 6% even when sample size is 100. Spearman-Rank and Kendall-Tau correlation coefficients are not much affected from the distribution shape where their type I error rates are quite similar to what was observed the type I error bivariate normal distribution. These two tests especially affected by the sample sizes, which is not surprising. Under these experimental conditions, the Winsorized and Permutation-based correlation coefficients, especially the Winsorized correlation coefficient, seem to be the most robust alternatives to the distribution shape, while all methods except Pearson correlation coefficient, seem to have similar the type I error rates as sample sizes are increased. When both the type I error rate and power were evaluated together, the Winsorized and Permutation-based correlation coefficients seem to be more advantageous over the Pearson, Spearman-Rank and Kendall-Tau correlation coefficients as they are more robust to the shape of the underlying distribution and sample size.

We also studied sampling distributions of the estimates for different correlation coefficients to evaluate the relative frequency of the positive estimates which provide an indirect indication of the relative test power of

Table 1. Type I error rates and test powers when samples taken from bivariate normal distribution.

Sample size	Correlation approach	$\alpha(1-\beta)$				
		True correlation coefficients				
		0	0.2	0.4	0.6	0.8
5	Pearson	4.93	9.00	16.50	30.42	58.23
	Spearman-Rank	1.70	2.02	3.03	5.75	13.44
	Kendal-Tau	1.67	2.00	3.04	5.88	13.45
	Permutation	6.03	11.01	17.85	30.84	55.43
	Winsorized	4.96	9.21	16.63	30.37	58.46
10	Pearson	5.14	13.77	32.17	62.56	92.76
	Spearman-Rank	4.95	7.50	17.86	39.01	75.75
	Kendal-Tau	4.16	7.48	17.01	38.15	74.73
	Permutation	5.75	16.27	36.22	68.36	94.88
	Winsorized	5.10	13.87	31.66	60.59	90.92
20	Pearson	4.83	22.01	57.86	91.53	99.88
	Spearman-Rank	4.89	11.96	37.15	76.70	98.66
	Kendal-Tau	4.77	11.49	36.11	75.41	98.56
	Permutation	5.88	24.04	57.34	91.33	99.90
	Winsorized	4.95	19.91	50.79	85.77	99.47
30	Pearson	4.95	28.17	72.71	97.74	99.99
	Spearman-Rank	4.97	17.03	54.40	92.31	99.94
	Kendal-Tau	4.99	16.74	54.30	92.26	99.93
	Permutation	5.06	30.75	74.28	97.76	100.00
	Winsorized	5.00	25.85	67.18	96.00	99.98
50	Pearson	4.92	40.08	90.10	99.90	100.00
	Spearman-Rank	5.02	37.77	86.62	99.76	100.00
	Kendal-Tau	5.01	25.89	78.51	99.39	100.00
	Permutation	5.14	42.23	90.14	99.97	100.00
	Winsorized	4.85	59.32	98.85	100.00	100.00
100	Pearson	4.93	64.10	99.46	100.00	100.00
	Spearman-Rank	4.96	59.83	98.95	100.00	100.00
	Kendal-Tau	5.00	47.47	97.75	100.00	100.00
	Permutation	5.29	34.20	82.18	99.41	100.00
	Winsorized	4.91	59.32	98.85	100.00	100.00

Table 2. Type I error rates and test powers when samples taken from bivariate *t*-distribution with 10 d.f.

Sample size	Correlation approach	$\alpha(1-\beta)$				
		True population coefficients				
		0	**0.2**	**0.4**	**0.6**	**0.8**
5	Pearson	5.02	8.89	15.73	28.30	52.95
	Spearman-Rank	1.64	1.99	3.05	5.39	11.85
	Kendal-Tau	1.64	2.01	3.15	5.84	13.43
	Permutation	5.94	10.09	17.94	29.06	51.48
	Winsorized	4.98	8.95	15.91	28.63	53.16
10	Pearson	5.14	13.15	30.29	58.24	89.65
	Spearman-Rank	4.94	7.47	17.00	37.02	72.21
	Kendal-Tau	4.69	7.33	16.99	38.25	75.05
	Permutation	5.08	13.92	34.11	64.31	92.07
	Winsorized	5.13	13.66	30.64	58.84	89.73
20	Pearson	4.94	19.97	52.14	87.01	99.64
	Spearman-Rank	4.94	11.74	35.56	74.01	98.16
	Kendal-Tau	4.75	11.66	36.12	75.44	98.52
	Permutation	4.91	20.83	51.68	87.72	99.59
	Winsorized	5.08	19.22	49.29	84.45	99.31
30	Pearson	4.98	26.35	68.06	96.47	99.99
	Spearman-Rank	5.06	16.16	52.67	90.88	99.92
	Kendal-Tau	4.92	16.42	54.38	92.16	99.94
	Permutation	4.81	26.35	68.56	97.03	100.00
	Winsorized	5.16	25.23	65.62	95.31	99.97
50	Pearson	4.99	37.27	86.92	99.77	100.00
	Spearman-Rank	5.03	36.29	85.33	99.67	100.00
	Kendal-Tau	4.93	25.94	78.53	99.41	100.00
	Permutation	5.33	38.29	87.26	99.71	100.00
	Winsorized	5.00	35.86	85.23	99.65	100.00
100	Pearson	4.91	59.65	98.93	100.00	100.00
	Spearman-Rank	4.99	58.38	98.71	99.99	100.00
	Kendal-Tau	5.01	47.31	97.77	99.99	100.00
	Permutation	5.32	59.69	98.82	100.00	100.00
	Winsorized	5.08	57.95	98.55	99.99	100.00

Table 3. Type I error rates and test powers when samples taken from bivariate log-normal distribution.

Sample size	Correlation approach	$\alpha(1-\beta)$				
		True population coefficients				
		0	**0.2**	**0.4**	**0.6**	**0.8**
5	Pearson	8.05	13.33	21.52	33.38	33.14
	Spearman-Rank	1.69	2.03	3.15	5.95	13.26
	Kendal-Tau	1.64	2.01	3.15	5.84	13.43
	Permutation	5.96	10.05	17.05	27.25	49.35
	Winsorized	7.81	13.42	21.12	33.34	54.59
10	Pearson	7.91	16.16	29.91	50.86	51.31
	Spearman-Rank	4.99	7.62	17.06	39.06	75.73
	Kendal-Tau	4.69	7.33	16.99	38.25	75.50
	Permutation	5.23	12.71	24.92	46.57	80.06
	Winsorized	5.47	13.37	28.64	54.37	87.05
20	Pearson	7.22	19.47	41.93	72.97	73.25
	Spearman-Rank	4.84	11.95	37.46	76.65	98.72
	Kendal-Tau	4.75	11.66	36.12	75.44	98.52
	Permutation	5.44	14.65	33.72	63.22	94.05
	Winsorized	5.44	18.67	45.86	80.05	98.97
30	Pearson	7.92	21.98	49.86	83.87	99.64
	Spearman-Rank	5.01	16.75	54.67	92.39	99.95
	Kendal-Tau	4.92	16.42	54.38	92.16	99.94
	Permutation	5.20	18.13	42.86	77.31	98.85
	Winsorized	5.37	23.89	60.43	92.98	99.95
50	Pearson	6.60	27.11	64.19	94.89	95.05
	Spearman-Rank	5.06	37.33	86.85	99.76	100.00
	Kendal-Tau	4.93	25.94	78.53	99.41	100.00
	Permutation	5.31	22.91	57.68	91.65	99.95
	Winsorized	5.37	32.89	80.31	99.24	100.00
100	Pearson	6.34	37.49	85.16	99.73	99.73
	Spearman-Rank	5.00	60.25	98.95	100.00	100.00
	Kendal-Tau	5.01	47.31	97.77	99.99	100.00
	Permutation	5.29	34.20	80.58	99.41	100.00
	Winsorized	5.16	52.50	97.20	99.99	100.00

these correlation coefficients. It is expected to observe a higher frequency of positive estimates for the correlation coefficient which corresponds to higher power when compared to the others. As expected, the number of positive correlation coefficient estimates increase as the sample size increases (**Figures 1-3**). In general, the test power of Permutation-based, Winsorized and Pearson correlation coefficients are higher than those of the Spearman rank and Kendal Tau correlation coefficients. There are significant differences between these tests especially when the sample size is small.

Figures 1-3 (also these) are examples of the sampling distributions of the five correlation coefficient methods for the true correlation coefficient of 0.6 and sample sizes of 5, 20 and 100, respectively. Each sampling distribution
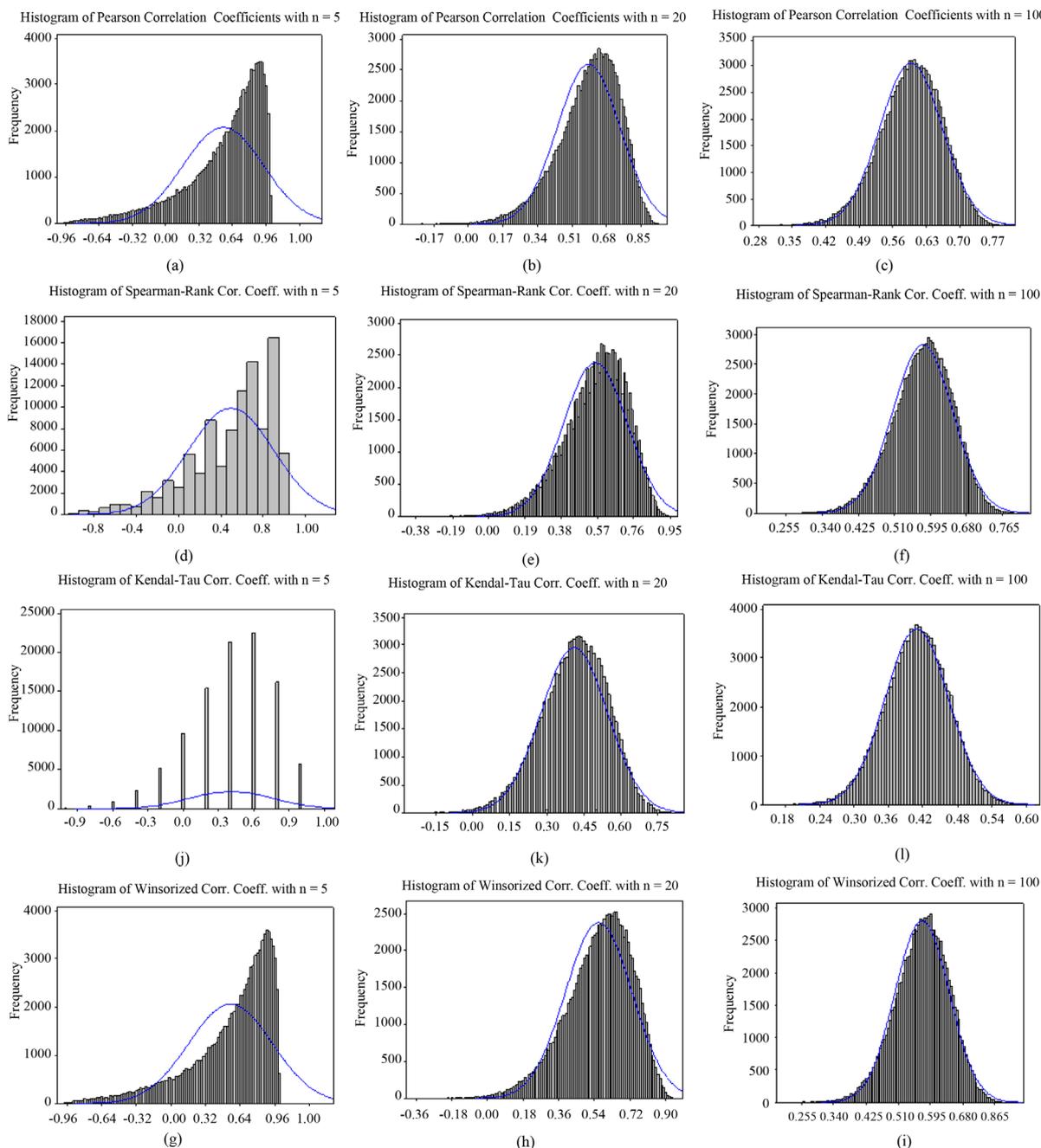
Figure 1. Sampling distributions of Pearson-Moment, Spearman-Rank, Kendal Tau and Winsorized correlation coefficients when samples were taken from bivariate normal distribution with $\rho = 0.60$.
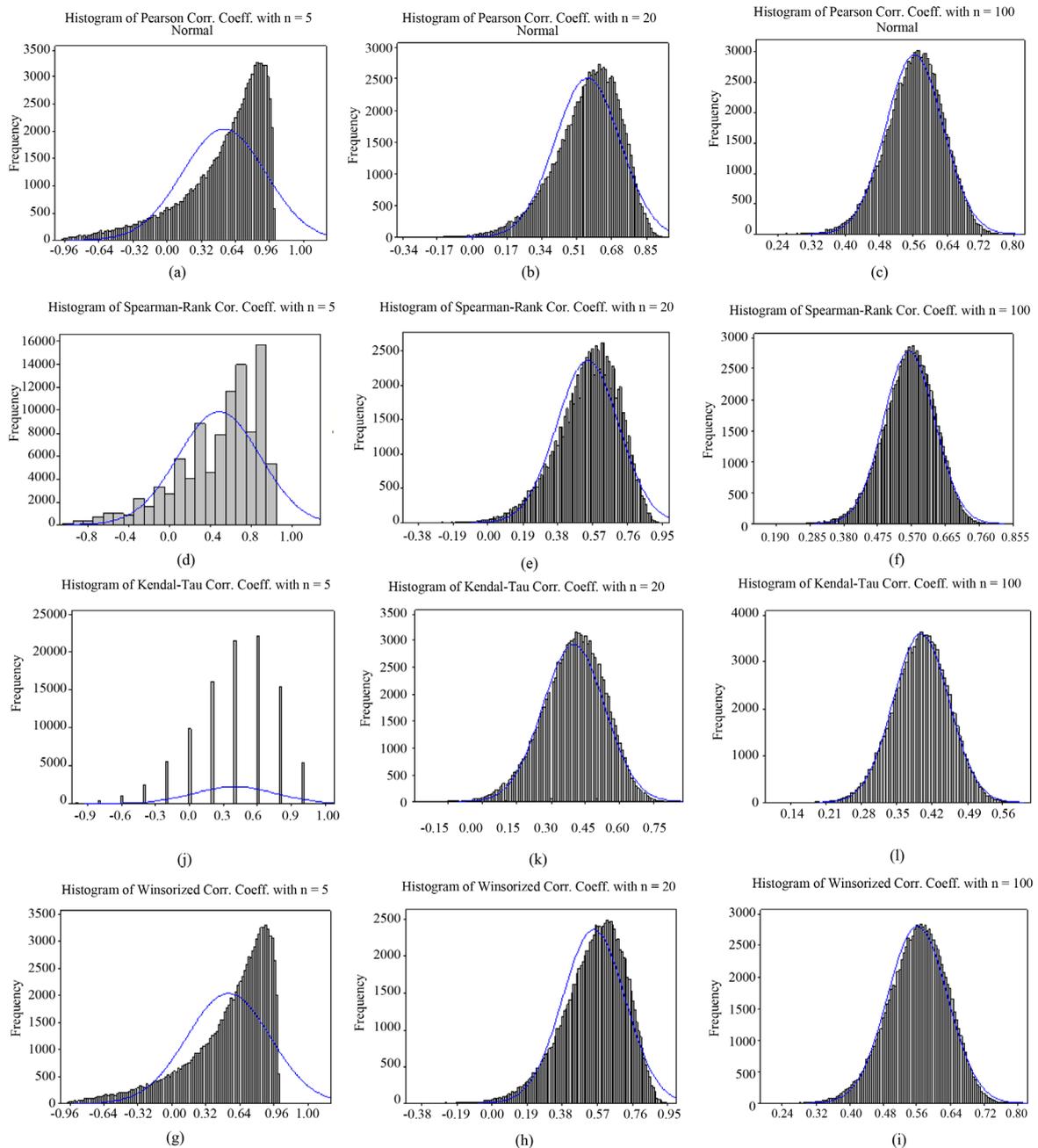
Figure 2. Sampling distributions of Pearson-Moment, Spearman-Rank, Kendal Tau and Winsorized correlation coefficients when samples were taken from bivariate *t*-distribution with 10 d.f.

is obtained from 50,000 simulations. Despite the smaller range, the frequency of positive estimates of Kendall's correlation coefficient is the least powerful of the five approaches.

4. A Real-Life Application

To illustrate the application of the five correlation coefficients discussed above. We used a data set that is obtained from an experiment to investigate the effects of three different feeding programs (AD: Adlibitum, R20: Feed Restriction Program 1, NF6: Feed Restriction Program 2) on growth curves of male chickens (**Table 4**), where the association between the back weight and the chest weight was of interest to the investigator.
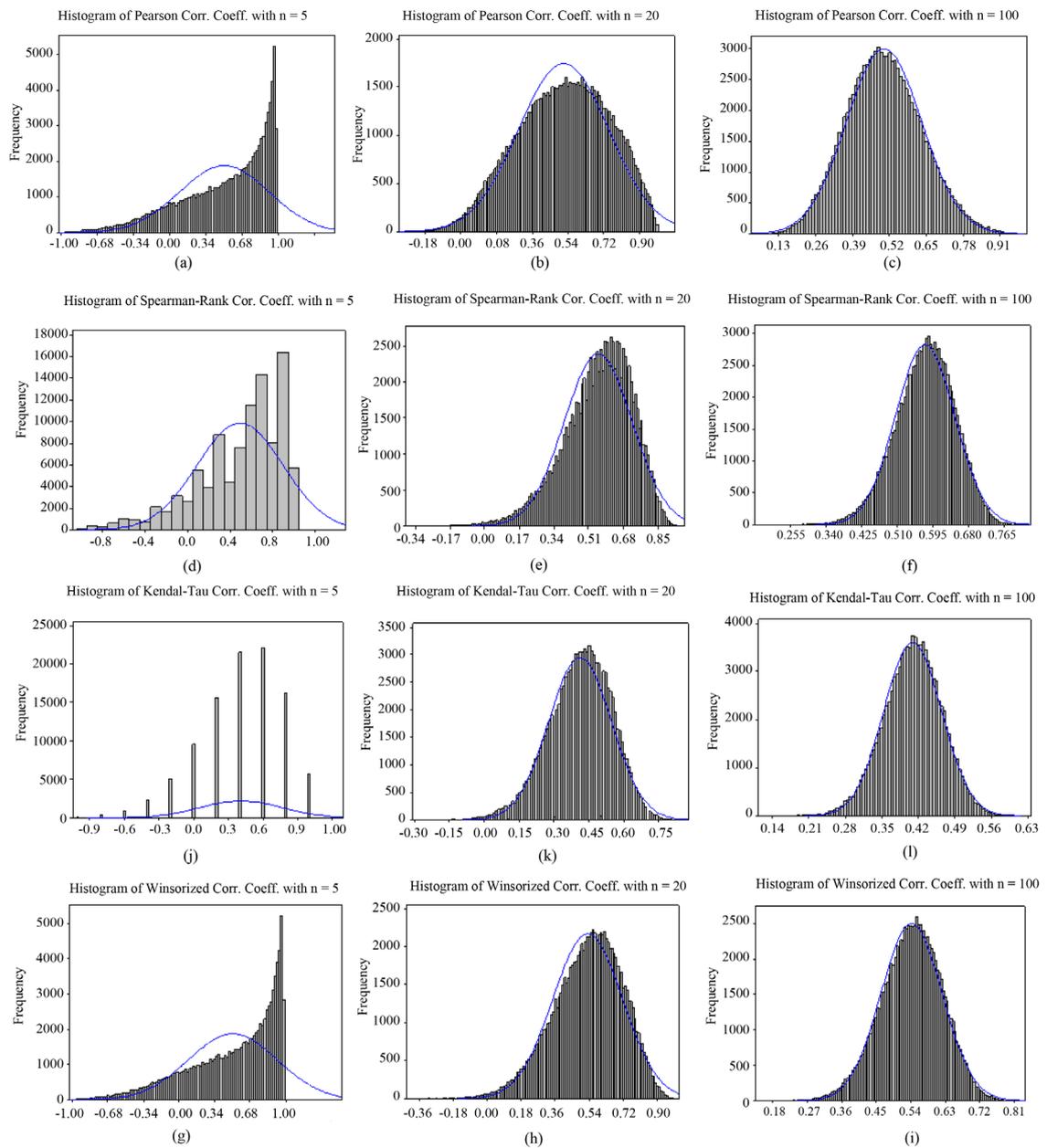
Figure 3. Sampling distributions of Pearson-Moment, Spearman-Rank, Kendal Tau and Winsorized correlation coefficients when samples were taken from bivariate log-normal distribution.

Table 4. Back weight (g) and chest weight (g) of chickens.

Back weight	Chest weight
365.0	479.5
381.0	401.5
256.0	346.5
367.5	403.0
350.5	380.0
257.5	334.0
402.0	377.5

4.1. Pearson Correlation Coefficient

$$r_{xy} = \frac{\sum d_x d_y}{\sqrt{\sum d_x^2 \sum d_y^2}} = 0.621 = 62.1\%$$

$$S_r = \sqrt{\frac{1-r_{xy}^2}{n-2}} = \sqrt{\frac{1-0.621^2}{5}} = 0.3505 \quad \text{and}$$

$$t = \frac{0.621}{0.3505} = 1.771.$$

4.2. Spearman Rank Correlation

Back weight (X)	Chest weight (Y)	Rank for X	Rank for Y	D ($X_i - Y_i$)	D^2
365.0	479.5	4	7	−3	9
381.0	401.5	6	5	1	1
256.0	346.5	1	2	−1	1
367.5	403.0	5	6	−1	1
350.5	380.0	3	4	−1	1
257.5	334.0	2	1	1	1
402.0	377.5	7	3	4	16

$$\sum d_i^2 = 30$$

$$r_s = 6 - \frac{6\sum d_i^2}{n(n^2-1)} = 6 - \frac{6(30)}{7(48)} = 0.464 = 46.4\%.$$

4.3. Kendall Tau Correlation

For this data set, the number concordant and discordant pairs are found to be 13 and 8, respectively. Thus, the Kendall Tau correlation coefficient and its significance test are calculated as follows:

$$r_{tau} = (C-D)/(n(n-1)/2) = (13-8)/21 = 0.238 = 23.8\%$$

$$Z = \frac{3*r_{tau}\sqrt{n(n-1)}}{\sqrt{2(2n+5)}} = \frac{3(0.238)\sqrt{42}}{\sqrt{38}} = 0.751.$$

4.4. Winsorized Correlation

Firstly, the data are sorted from the smallest to the largest in each cell. Then, the smallest and the largest values are replaced with the observations closest to them. Let Winsorized percentage, g, be 10.0%. In this case, the new data set would be as follows:

Back weight (X)	Chest weight (Y)
365.0	**403.0**
381.0	401.5
257.5	346.5
367.5	403.0
350.5	380.0
257.5	**346.5**
381.0	377.5

$$\sum d_x d_y = (365.0) \times (403.0) + (381.0) \times (401.5) + \cdots + (381.0) \times (377.5) - \frac{(2360) \times (2658)}{7} = 7508.286$$

$$\sum d_x^2 = 365.0^2 + 381.0^2 + \cdots + 381.0^2 - \frac{(2360)^2}{7} = 18408.857$$

$$\sum d_y^2 = 403.0^2 + 401.5^2 + \cdots + 377.5^2 - \frac{(2658)^2}{7} = 3770.429$$

$$r_w = \frac{\sum d_x d_y}{\sqrt{\sum d_x^2 \sum d_y^2}} = \frac{7508.286}{\sqrt{(18408.857)(3770.429)}} = 0.901 = 90.1\%$$

$$S_{rw} = \sqrt{\frac{(1 - rw_{xy}^2)}{(n-2)}} = \sqrt{\frac{(1 - 0.901^2)}{5}} = 0.194$$

$$T_W = \frac{r_w}{S_{rw}} = \frac{0.901}{0.194} = 4.64.$$

4.5. Permutation-Based Correlation Coefficient

For this data set, it is possible to form $7! = 7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1 = 5040$ permutation samples. From these possible permutations, we have computed the permutation-based *p*-value as 0.152. Correlation coefficient values and hypotheses testing about these correlation coefficients are given in **Table 5**. As seen in **Table 5**, while the Winsorized correlation coefficient rejected the null hypothesis, the Pearson, Permutation-based, Spearman and Kendal-Tau correlations accepted the null hypothesis. Therefore, due to differences in their theory, computation steps, and assumptions it is not surprise to get different results even the same data set is used.

5. Discussion

In any research area, association between two variables is often of interest to the researcher; for example, the researcher may be interested in estimating the association between weight and systolic blood pressure, or plant length and yield, etc. In doing so, the researcher wishes to test for the existence of an association as well as its strength and direction if one exists. To do that, among various available methods, the researcher needs to choose the one that is most suitable to the data at hand, which may not be an easy task. Although most commonly employed in practice, many studies have shown that the Pearson-moment correlation coefficient is not robust to the violations of its assumptions, which prompts the researcher to consider alternatives such as Spearman-Rank, Kendal-Tau, and Pearson correlation-based nonlinear transformation, Permutation-based and Winsorized correlation coefficients, etc., where Spearman-Rank and Kendal-Tau correlation coefficients are the most cited alternatives. However, although the advantages of these two correlation coefficients for the ordinal and ranked data are emphasized, their performance under continuous data is not clear. In addition, the performance of the Spearman-Rank correlation coefficient has not been investigated thoroughly under many experimental conditions which the researcher faces in practice. Another practical approach is to normalize the data via a suitable transformation so to use Pearson-correlation coefficient [15] [16] [28] [29]. However, nonlinear transformations

Table 5. Correlation coefficients, the corresponding critical values, and decision about null hypothesis.

Correlation coefficient	Values	Critical table value	Decision
Pearson	0.621	2.571	H_0: accept
Spearman	0.464	2.571	H_0: accept
Kendall	0.238	1.96	H_0: accept
Winsorized	**0.901**	**2.571**	**H_0: reject**
Permutation	0.152	0.05	H_0: accept

are not given satisfactory results especially when distributions are highly skewed and sample size is small. Furthermore, nonlinear transformations can cause to interpretational ambiguity [13] [14]. Thus, it necessitates using other robust alternatives such as Winsorized and Permutation-based correlation coefficients when assumptions of Pearson-moment correlation coefficient are violated. However, there has been a need to extensively investigate the performance of these two correlation coefficients under different experimental conditions commonly met in practice.

Simulation results showed that unless severe deviations from the normality assumptionexist, Pearson correlation coefficient is a method of choice, which is the case under both bivariate normal distribution and bivariate *t*-distribution with 10 d.f, where there is no skewness. However, it is sensitive to excessive deviations from the normality especially to skewness with small sample sizes (n < 50).

Although Spearman-Rank correlation coefficient is the most recommended when assumptions of Pearson-moment correlation coefficient is not met, in our simulations, its performance as well as the performance of Kendal-Tau are quite poor. Our findings seem to contradict with those of Fowler [10] who reported that Spearman's correlation was more powerful than Pearson's r across a range of non-normal bivariate distributions. Likewise, Zimmerman and Zumbo [11] reported that Spearman's correlation was more powerful under mixed-normal and non-normal distributions. They also reported that when studying with exponential distributions, Spearman's correlation retained the type I error rates at or below the nominal level whereas Pearson's correlation produced inflated the type I error. Indeed, as stated in many studies, using the Spearman-Rank and Kendal Tau correlation coefficients are more appropriate when data are ordinal or ranked [24] [30]-[32]. The apparent difference between our results and those discussed above may come from the experimental conditions such as distribution shapes, sample size, number of simulations taken into considered (Note that we worked with 50,000 simulation samples per condition).

In general, Permutation-based Pearson correlation coefficient and especially the Winsorized correlation coefficient are most robust to the distribution shape, sample size, and outliers. Thus, it is possible to recommend both Permutation-based correlation coefficient and Winsorized correlation coefficient over the Spearman-Rank and Kendall-Tau correlation coefficients when assumptions of Pearson correlation coefficient are not satisfied. Hayes [2] reported that especially when normality assumption is violated, permutation tests tend to do well at controlling the type I error rate, which was similarly argued by Good [19] as well. Good [33] and Mielke and Berry [34] recommended using Permutation-based procedures, especially when sample sizes are small and the variables of interest are non-normally distributed, which are conclusions supported by our study as well. However, it must be kept in mind as suggested by Keller-McNulty and Higgins [35], Rasmussen [15], Hayes [2] that permutation tests do not always solve all assumption violation problems, and simulation results on this issue have been mixed.

6. Conclusion

To conclude, we recommend the use of either Winsorized correlation coefficient or Permutation-based Pearson correlation for investigating the relationship between two continuous variables of interest. Between the two, the permutation-based Pearson correlation is more computationally challenging, as it is critical to generate sufficiently large number of permutation samples to avoid misleading results [36]. For example, for n = 5, there exists 5! = 120 permutation samples, which can be generated very easily. For n = 8, we have 40,320, and for n = 10, we have 3,628,800 possible permutations. Any sample size beyond n = 10 may be too difficult to generate all possible permutations, which necessitates to sample from possible permutations, where such a sample must be sufficiently large. Although Hayes [2] [37] [38] proposed to limit the number of permutation samples to 5000, with today's computing power, a sample of as big as 100,000 or even bigger sample size, can be easily achieved.

References

[1] Edgell, S. and Noon, S. (1984) Effect of Violation of Normality on the t Test of the Correlation Coefficient. *Psychological Bulletin*, **95**, 576-583. http://dx.doi.org/10.1037/0033-2909.95.3.576

[2] Hayes, A.F. (1996) PERMUSTAT: Randomization Tests for the MacIntosh. *Behavior Research Methods*, *Instruments*, *& Computers*, **28**, 473-475. http://dx.doi.org/10.3758/BF03200530

[3] Bonett, D.G. and Wright, T.A. (2000) Sample Size Requirements for Estimating Pearson, Kendall and Spearman Correlations. *Psychometrics*, **65**, 23-28. http://dx.doi.org/10.1007/BF02294183

[4] Zimmerman, D.W., Zumbo, B.D. and Williams, R.H. (2003) Bias in Estimation and Hypothesis Testing of Correlation. *Psychologica*, **24**, 133-158.

[5] Wilcox, R.R. (2001) Fundamentals of Modern Statistical Methods: Substantially Improving Power and Accuracy. Springer, New York. http://dx.doi.org/10.1007/978-1-4757-3522-2

[6] Warner, R. (2008) Applied Statistics: From Bivariate through Multivariate Techniques. Sage Publications, Inc., Thousand Oaks.

[7] Triola, M. (2010) Elementary Statistics. 11th Edition, Addison-Wesley/Pearson Education, Boston.

[8] Witte, R. and Witte, J. (2010) Statistics. 9th Edition, Wiley, New York.

[9] Blair, R. and Lawson, S. (1982) Another Look at the Robustness of the Product-Moment Correlation Coefficient to Population Non-Normality. *Florida Journal of Educational Research*, **24**, 11-15.

[10] Fowler, R. (1987) Power and Robustness in Product-Moment Correlation. *Applied Psychological Measurement*, **11**, 419-428. http://dx.doi.org/10.1177/014662168701100407

[11] Zimmerman, D. and Zumbo, B. (1993) Significance Testing of Correlation Using Scores, Ranks, and Modified Ranks. *Educational and Psychological Measurement*, **53**, 897-904. http://dx.doi.org/10.1177/0013164493053004003

[12] Bishara, A.J. and Hittner, J.B. (2012) Testing the Significance of a Correlation with Non-Normal Data: Comparison of Pearson, Spearman, Transformation, and Resampling Approaches. *Psychological Methods*, **17**, 399-417.

[13] Osborne, J. (2002) Notes on the Use of Data Transformations. *Practical Assessment*, *Research & Evaluation*, **8**. http://pareonline.net/getvn.asp?v=8&n=6

[14] Tabachnick, B. and Fidell, L. (2007) Using Multivariate Statistics. 5th Edition, Allyn & Bacon/Pearson Education, Boston.

[15] Rasmussen, J. (1989) Data Transformation, Type I Error Rate and Power. *British Journal of Mathematical and Statistical Psychology*, **42**, 203-213. http://dx.doi.org/10.1111/j.2044-8317.1989.tb00910.x

[16] Dunlap, W., Burke, M. and Greer, T. (1995) The Effect of Skew on the Magnitude of Product-Moment Correlations. *Journal of General Psychology*, **122**, 365-377. http://dx.doi.org/10.1080/00221309.1995.9921248

[17] Calkins, D.S. (1974) Some Effects of Non-Normal Distribution Shape on the Magnitude of the Pearson Product Moment Correlation Coefficient. *Interamerican Journal of Psychology*, **8**, 261-288.

[18] Wilcox, R.R. (1993) Some Results on a Winsorized Correlation Coefficient. *British Journal of Mathematical and Statistical Psychology*, **46**, 339-349. http://dx.doi.org/10.1111/j.2044-8317.1993.tb01020.x

[19] Good, P. (2009) Robustness of Pearson Correlation. *Interstat*, **15**, 1-6.

[20] Kotz, S., Balakrishnan, N. and Johnson, N.L. (2000) Continuous Multivariate Distributions: Volume 1: Models and Applications. 2nd Edition, John Wiley & Sons, Inc., New York. http://dx.doi.org/10.1002/0471722065

[21] Ruscio, J. and Kaczetow, W. (2008) Simulating Multivariate Non-Normal Data Using an Iterative Algorithm. *Multivariate Behavioral Research*, **43**, 335-381. http://dx.doi.org/10.1080/00273170802285693

[22] Balakrishnan, N. and Lai, C.D. (2009) Continuous Bivariate Distributions. 2nd Edition, Springer, New York.

[23] Chok, N.S. (2010) Pearson's versus Spearman's and Kendall's Correlation Coefficients for Continuous Data. Master's Thesis, University of Pittsburgh, Pittsburgh.

[24] Sheskin, D.J. (2007) Handbook of Parametric and Nonparametric Statistical Procedures. 4th Edition, Chapman & Hall/CRC, Boca Raton.

[25] Tanizaki, H. (2004) On Small Sample Properties of Permutation Tests: Independence between Two Samples. *International Journal of Pure and Applied Mathematics*, **13**, 235-243.

[26] Ernst, M.D. (2004) Permutation Methods: A Basis for Exact Inference. *Statistical Science*, **19**, 676-685. http://dx.doi.org/10.1214/088342304000000396

[27] Knight, W. (1966) A Computer Method for Calculating Kendall's Tau with Ungrouped Data. *Journal of the American Statistical Association*, **61**, 436-439. http://dx.doi.org/10.1080/01621459.1966.10480879

[28] Kowalski, C.J. and Tarter, M.E. (1969) Co-Ordinate Transformations to Normality and the Power of Normal Tests for Independence. *Biometrika*, **56**, 139-148. http://dx.doi.org/10.1093/biomet/56.1.139

[29] Rasmussen, J. and Dunlap, W. (1991) Dealing with Non-Normal Data: Parametric Analysis of Transformed Data vs. Nonparametric Analysis. *Educational and Psychological Measurement*, **51**, 809-820. http://dx.doi.org/10.1177/001316449105100402

[30] Siegel, S. (1957) Nonparametric Statistics. *The American Statistician*, **11**, 13-19.

[31] Lieberson, S. (1964) Limitations in the Application of Non-Parametric Coefficients of Correlation. *American Sociological Review*, **29**, 744-746. http://dx.doi.org/10.2307/2091428

[32] Khamis, H. (2008) Measures of Association: How to Choose? *Journal of Diagnostic Medical Sonography*, **24**, 155-162. http://dx.doi.org/10.1177/8756479308317006

[33] Good, P. (2005) Permutation, Parametric and Bootstrap Tests of Hypotheses. 3rd Edition, Springer-Verlag, New York.

[34] Mielke, P.W. and Berry, K.J. (2007) Permutation Methods. A Distance Function Approach. 2nd Edition, Springer, New York.

[35] Keller-McNulty, S. and Higgins, J.J. (1987) Effect of Tail Weight and Outliers on Power and Type-I Error of Robust Permutation Tests for Location. *Communications in Statistics B-Simulation and Computation*, **16**, 17-35. http://dx.doi.org/10.1080/03610918708812575

[36] Akkartal, E., Mendes, M. and Mendes, E. (2010) Determination of Suitable Permutation Numbers in Comparing Independent Group Means: A Monte Carlo Simulation Study. *Journal of Scientific & Industrial Research*, **69**, 422-425.

[37] Hayes, A.F. (1998) SPSS Procedures for Approximate Randomization Tests. *Behavior Research Methods*, *Instruments*, *& Computers*, **30**, 536-543. http://dx.doi.org/10.3758/BF03200687

[38] Hayes, A.F. (2000) Randomization Tests and the Equality of Variance Assumption When Comparing Group Means. *Animal Behaviour*, **59**, 653-656. http://dx.doi.org/10.1006/anbe.1999.1366