EÜFBED - Fen Bilimleri Enstitüsü Dergisi Cilt-Sayı: 7-1 Yıl: 2014 *133-148*

A NEW ROBUST APPROACH TO TESTING SIGNIFICANCE OF CORRELATION

KORELASYONUN ANLAMLILIK TESTİNE YENİ BİR SAĞLAM YAKLAŞIM

Mehmet KARAHASAN[1]

Muğla Sıtkı Koçman Üniversitesi, Fen Fakültesi, İstatistik Bölümü, Kötekli, Muğla

Geliş Tarihi: 13 Aralık 2013 **Kabul Tarihi:** 17 Nisan 2014

ÖZET

Bu çalışmada iki normal dağılımlı rasgele değişken arasındaki korelâsyonun anlamlılık testi için yeni bir test yaklaşımı önerilmektedir. Bu yeni yaklaşım korelâsyonun işareti ile uyumlu olmayan aykırı değerlere karşı sağlamdır. Korelasyonun anlamlılığının testi için kullanılan geleneksel Pearson t-testi ile önerilen test yaklaşımının güç performanslarını karşılaştırmak amacıyla, bu tür aykırı değerlerin birkaç tanesinin varlığı durumunda bir benzetim çalışması gerçekleştirilmiştir. Bu benzetim çalışması sonucunda, önerilen yaklaşımın korelâsyonlar için olan Pearson t-testinden daha iyi performans gösterdiği görülmüştür. Bu tür aykırı değerler örneklemin yaklaşık olarak %5 ya da %6'sından daha fazlasını oluşturmamak kaydıyla yeni yaklaşıma ilişkin bu iyi performansın devam ettiği söylenebilmektedir.

Anahtar kelimeler: Kolmogorov-Smirnov uyum iyiliği testi, Cramer-von Mises uyum iyiliği testi, Deneysel güç karşılaştırması, Simülasyon

ABSTRACT

In this study, a new testing approach to significance test for correlation between two normally distributed random variables is proposed. This new approach is robust to the outliers that are incompatible with the sign of the correlation. To evaluate and compare power performance of the proposed approach with that of usual the Pearson t-test for correlation in the case of this type of a few outliers, a simulation study is conducted. As a result of the simulation study, it has been found that the proposed approach performs better than the usual t-test in the presence of such outliers. In addition, it seemed that the nice performance of the new approach continues to exist provided that thesde outliers constitute approximately no more than 5% or 6% of the sample.

Keywords: Kolmogorov-Smirnov goodness of fit test, Cramer-von Mises goodness of fit test, Empirical power comparison, Simulation

***Sorumlu Yazar: mkarahas@mu.edu.tr**

1. INTRODUCTION

In research activities of many disciplines, testing statistically significance of correlation between two continuous variables based on random samples is frequently concerned. Researchers generally search for whether there exists a linear relationship between the variables of interest. In this context, the t-test based on Pearson's Product Moment Correlation Coefficient is one alternative to use in the case of bivariate normal distribution (Wackerly, 2007):

$$t = (r\sqrt{n-2} / \sqrt{1-r^2}) \sim t_{n-2} \quad (1)$$

where r is the sample correlation coefficient:

$$r = \sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}) \Big/ \sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2}$$

The test statistics in (1) is shown to be distributed as t-distribution with degrees of freedom n-2. With the test statistics in (1), the null hypothesis $H_0 : \rho = 0$ is tested versus the alternative hypothesis $H_A : \rho \neq 0$. In the case of bivariate normal distribution, the null hypothesis corresponds to independence or equivalently no linear relationship between the variables of interest (Wackerly, 2007).

However, King (2003) reports that "r is not as robust as is commonly asserted, especially when the bivariate surface is non-normal and dependence exists" according to his review of the literature about the robustness of r. Based on the review, the author also states that this non-robustness has reflections even in terms of statistical significance testing. Similarly, Huber (2004) indicates non-robustness of r and shift of its value anywhere in the interval (-1, 1) due to one single, sufficiently bad outlying pair. In this context, Wilcox (2012) states that "r is not resistant; a single unusual point can dominate its value". Pernett et al. (2013) supports this argument stating that "r is overly sensitive to outliers. Indeed, a single outlier can result in a highly inaccurate summary of the data". Further, El-Fallah and El-Salam (2006) show high sensitivity of Pearson's correlation coefficient to the presence of outliers. A few outliers can change significantly the value of r because the sample product moment correlation estimator is not a very robust estimator to outliers in data, which are unusual observations compared to majority of the data (Evandt et al., 2004). This is likely the case when the outliers are incompatible with the sign of the current correlation present in data.

Such limitation of r may affect the power of the test in (1) negatively, i.e., the probability of rejecting the null hypothesis of no correlation gets lower when there is indeed meaningful correlation between the variables of interest. Bishara and Hittner (2012) take notice of the situation that the Pearson t-test results in inflated Type I errors as well as low power compared to alternative methods for highly kurtotic distributions in their simulation study. In the light of these finding, they suggest that robust alternatives to the Pearson t-test should be used when distributions are highly kurtotic and, thus especially prone to outliers on one or both tails.

Outliers can occur because of many reasons such as randomness, heavy tailed population distributions, and measurement or recording errors, and mixture of two distributions. Outliers can reduce or distort performance of statistical methods. In order to alleviate the effect of outliers on statistical inference, one alternative is to use robust methods, which are not affected unduly by outliers.

In order to overcome the likely low power property of the t-based on r in the case of particular outliers in data, a new robust approach to test significance of correlation between variables of interests is proposed in this paper. Then a simulation study is conducted to compare power performance of the t-test based on r and the new testing approach in the case of the particular outliers.

2. METHODS

In this section, a new approach to test significance of correlation will be proposed in the case that the outliers are in incompatible with direction of the current correlation present in data: when there is a positive correlation, the outlier points (X, Y) behave as if there were negative correlation or when there is a negative correlation, the outlier points (X, Y) act as if there were positive correlation. An example of this situation is depicted in Figure 1.

To this end, in this section, first, the rationale behind the new robust significance testing approach will be explained for the correlation coefficient ρ. Then, the algorithm for the new approach will be given. Finally, simulation settings for empirical power comparisons are described.

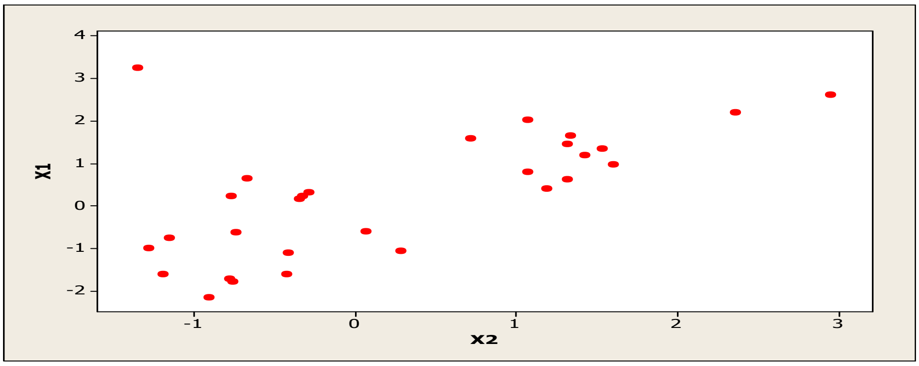


Figure 1. Scatter plot of a simulated random sample of 30 pairs of observations from bivariate normal distribution with parameters $\mu_X = 0$, $\mu_Y = 0$, $\sigma_X = 1$, $\sigma_Y = 1$ and $\rho = 0.7$ mixed 1 outlier.

2.1 The underlying Ideas

Assume that (X_i, Y_i), $i = 1,2,\ldots,n$ are a random sample from a bivariate normal distribution with the parameters μ_X, μ_Y, σ_X, σ_Y, and ρ. When performing significance testing with the t-test in (1), the pairs of observations (X_i, Y_i) of which at least one element is unusually large or small, namely outliers, can distort the value of the test statistics. This is the case since the whole sample is projected onto a single statistics. Thus, it is expected that the ability of making right decisions with the test decreases, i.e. the probability of rejecting the hypothesis of no correlation with the test when there exists non zero correlation decreases for the test statistics.

In order to overcome this problem, the entire sample should be used in statistical testing without converting it into a single statistics that is likely to be unduly influenced by the outliers. Thereby, the effect of the outliers can be reduced. In turn, the testing procedure using the entire sample will not be as powerful as the t-test in (1) when the all assumptions hold with no outliers present.

The sample of the pairs (X_i, Y_i), $i = 1,2,\ldots,n$ is two dimensioned. It would be convenient if it is represented in one dimension, i.e., only in observations of one variable. However, which transformation to use is very important; the transformed observations must contain all the necessary information about the property being tested. Since no correlation corresponds to independence for the bivariate normal distribution, the product of observations $X_i Y_i$ will have necessary information about the independence. In order to free the variables from their measurement scale, the observations are standardized, then multiplied with each other: $Z_{X_i} Z_{Y_i}$, $i = 1,2,\ldots,n$.

2.2 The Algorithm for the New Proposed Significance Testing Approach

The application of the idea presented in the Subsection 2.1 can be expressed in steps as follows.

1. A random sample of the observation pairs (X_i, Y_i), $i = 1,2,\ldots,n$ is obtained from a bivariate distribution that is assumed to be bivariate normal.
2. Standardizing the pairs of observations (X_i, Y_i), the standardized pairs of observations (Z_{Xi}, Z_{Yi}), $i = 1,2,\ldots,n$ are obtained. If X and Y independent random variables, Z_{Xi} and Z_{Yi} will be independent and distributed as standard normal.
3. By using a goodness of fit test, it is tested whether the sample of $Z_{Xi} Z_{Yi}$, $i = 1,2,\ldots,n$ comes from the product-normal distribution that is the distribution of the product $Z_X Z_Y$ of independent and standard normally distributed random variables Z_{X_i} and Z_{Y_i}.
4. If it is not rejected that the sample of $Z_{Xi} Z_{Yi}$, $i = 1,2,\ldots,n$ comes from the product-normal distribution, it is concluded that there is no correlation between X and Y random variables, meaning that the variables are independent or that there is no linear relationship between these two variables. Otherwise, it is concluded that the random variables X and Y are correlated with each other.

Since the new approach involves using a goodness of fit test, its power is not expected to be as good as the t-test based on r in (1) when the assumptions of the t-test are met with no outliers in data. On the other hand, when some outliers, especially the ones incompatible with the sign of correlation are present in data, the power of the new testing approach is expected to be better than the t-test in (1). With the judicious choice of the goodness of fit test to be employed, this power can be improved further.

In this study two goodness of fit test, namely Kolmogorov-Smirnov test and Cramer von-Mises test are considered. First, the formula for the Kolmogorov-Smirnov test is as follows (Bain and Engelhardt, 1992).

$$D = \max(D^+, D^-)$$

where $x_{i:n}$ is the observed value of ith order statistics in a random sample of size n and D^+, D^- respectively defined as

$$D^+ = \max_i \left(\frac{i}{n} - F(x_{i:n}) \right) \quad \text{and} \quad D^- = \max_i \left(F(x_{i:n}) - \frac{(i-1)}{n} \right).$$

The hypotheses tested by the Kolmogorov-Smirnov test is stated as

H_0 : The sample comes from the distribution with cumulative distribution function $F(x)$

H_1 : The sample does not come from the distribution with cumulative distribution function $F(x)$

where $F(x)$ denotes the completely specified distribution function for the population being considered. The distribution of these test statistics Ds do not depend on F, that is, statistics Ds are distribution-free. Furthermore, asymptotic critical values of these statistics were derived by Stephens (1974, 1977) and some modifications of these values are made to take into account small sample sizes n (Bain and Engelhardt, 1992). The tables of critical values for $n \leq 40$ associated with the Kolmogorov-Smirnov test is given in Table A14 of Conover (1999).

Second, the test statistics associated with the Cramer von-Mises test is given as follows (Bain and Engelhardt, 1992).

$$CM = \frac{1}{12n} + \sum_{i=1}^{n}\left(F(x_{i:n}) - \frac{(i-0.5)}{n} \right)^2$$

where $x_{i:n}$ and F(x) are defined as in the Kolmogorov-Smirnov test. The same hypotheses H_0 and H_1 above are tested by Cramer- Von Mises test. Scott (2000) gave tables of statistics CM and compared them with the limiting distribution.

2.3. Simulation Settings

In order to compare power performances of the t-test in (1) and the new testing approach proposed in this paper, random samples of size n=10, 20, 30, 40 and 50 have been generated from the bivariate normal distribution with the parameters $\mu_X = 0$, $\mu_Y = 0$, $\sigma_X = 1$, $\sigma_Y = 1$, and ρ. The values of ρ are varied from 0 to 0.90 in the increments of 0.10. Negative values of correlation are not considered since the results parallel to those of positive ones are expected.

The outliers incompatible with the sign of correlation are included in the data: the simulated samples are mixed with 1 or 2 such outlier points. The reason for choosing 1 or 2 outliers is to see the effect of a few outliers of this type on power of the t-test in (1). The outliers are generated from the bivariate distribution with the means $\mu_X = 2.5$, $\mu_Y = -2.5$, while the standard deviations and correlation for the outlier distribution are taken the same as those of the distribution from which the data are generated.

For each combination of n, ρ and mix ratio value, 1000 samples mixed with the particular outliers have been generated, then significance correlation tests applied at the level of $\alpha = 0.05$ and the results of these tests recorded. In other words, for each combination of n, ρ and mix ratio value, 1000 repetitions have been done to obtain empirical power estimates, which is the number of rejections of the null hypothesis of no correlation divided by the number of repetitions. When $\rho = 0$, the power estimates turn into estimates of probability of Type I error, which is the probability of rejecting the null hypothesis when the null hypothesis is true. Simulations have been performed by a computer program of Java codes developed specifically for this study. The graphs in figures have been obtained from Minitab 14 statistical software package. As to precision of rates of rejections, with 1000 simulations, the largest possible standard error for the proportion of rejections of no correlation is

$$\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{0.5(1-0.5)}{1000}} = 0.0158 .$$

Therefore, an approximate 95% confidence interval for the proportion of rejections is ± 2 standard errors, that is, $\pm 3\%$.

3. FINDINGS

Table 1 through Table 5 give empirical Type I error probabilities and power estimates of the correlation t-test and the new testing approach denoted by O-KS and O-CVM using Kolmogorov-Smirnov (KS) and Cramer Von-Mises (CVM) tests respectively for sample sizes of n=10, 20, 30, 40, and 50 with 1 and 2 outliers mixed. Also, Type I error probability and power performances for these methods of testing are displayed in Figure1 through Figure 5.

Empirical Type I error probability of the t-test is considerably much larger than $\alpha = 0.05$ and that of the new testing approach for all sample sizes. However, the probability of Type I error for the new testing approach tends to be not much larger than $\alpha = 0.05$. In addition, it is seen that as the number of outliers increases, Type I error probability of the new testing approach proposed in this paper becomes larger.

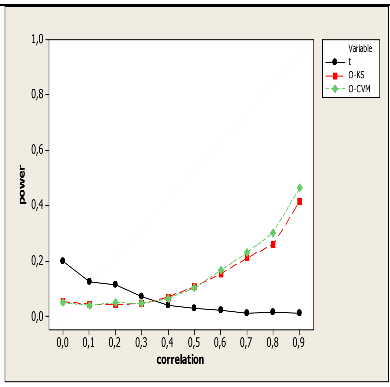
As for the power performances, it seems that none of the tests has acceptable power level for n=10. All the tests considered in this paper appear to have generally low power levels when the level of correlation is lower than 0.50 for sample size 20 or above. For the range of correlation values from 0.50 to 1, it is observed that the new testing approach is more powerful than the correlation t-test. This superior performance is more noticeable in the case of samples with 2 outliers.

One thing to notice is that the performance of the new testing approach does not differ considerably according to use of Kolmogorov-Smirnov test or Cramer-Von Mises test in terms of power and probability of Type I error.

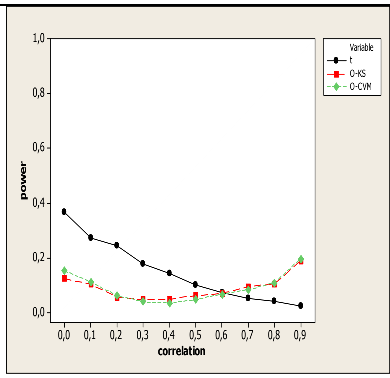
Table 1. The empirical power estimates of significance tests of correlation (1000 repetitions)

n = 10	mixed with 1 outlier			mixed with 2 outliers		
ρ *values*	*t-test*	*O-KS*	*O-CVM*	*t-test*	*O-KS*	*O-CVM*
0.0	0.199	0.054	0.050	0.368	0.124	0.153
0.1	0.126	0.044	0.039	0.274	0.105	0.112
0.2	0.113	0.043	0.050	0.243	0.057	0.062
0.3	0.072	0.046	0.045	0.179	0.047	0.04
0.4	0.041	0.069	0.064	0.144	0.049	0.036
0.5	0.030	0.108	0.102	0.101	0.062	0.048
0.6	0.023	0.154	0.166	0.073	0.071	0.067
0.7	0.012	0.212	0.230	0.052	0.093	0.084
0.8	0.015	0.260	0.303	0.041	0.106	0.109
0.9	0.010	0.414	0.465	0.023	0.190	0.197

*O-KS denotes the new test approach using KS test, while O-CVM denotes the new test approach using CVM test.



(a) n=10 with 1 outlier



(b) n=10 with 2 outliers

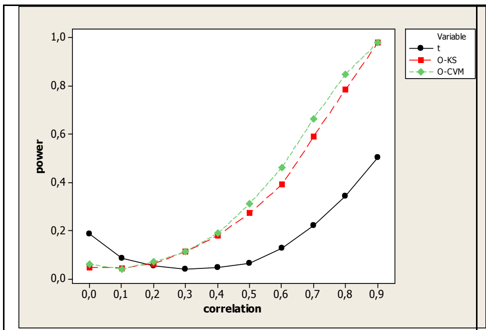
Figure 2. The scatter plots of empirical power estimates of significance tests of correlation

Table 1 and Figure 2 indicate considerably larger Type I error rate than the nominal rate of 0.05 for t-test in both cases whereas those of new test approach are not so large in comparison to 0.05. In addition, all of the methods have unacceptably low power for non zero correlation values. Furthermore, as correlation values increase, the power for t-test tends to decrease while the power values for the new test approach tend to increase.

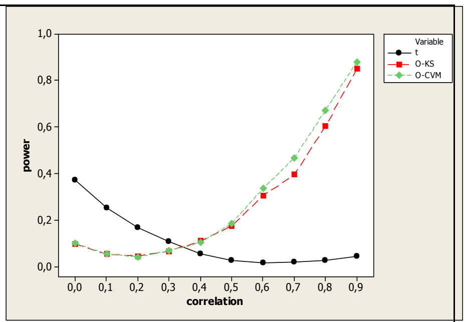
First, Table 2 and Figure 3 show that Type I error rate of t-test inflates as the number of outliers increases. Second, when the outliers increase in number as in the cases of samples mixed with 2 outliers, the power for t-test declines as correlation values become larger. In contrast, the power values for the new test approach become larger with larger correlation values.

Table 2. The empirical power estimates of significance tests of correlation (1000 repetitions)

n = 20	*mixed with 1 outlier*			*mixed with 2 outliers*		
ρ *values*	*t-test*	*O-KS*	*O-CVM*	*t-test*	*O-KS*	*O-CVM*
0	0.185	0.049	0.060	0.374	0.098	0.101
0.10	0.084	0.045	0.041	0.252	0.056	0.057
0.20	0.056	0.063	0.070	0.168	0.047	0.043
0.30	0.041	0.112	0.113	0.110	0.068	0.07
0.40	0.049	0.181	0.189	0.055	0.112	0.105
0.50	0.064	0.272	0.312	0.029	0.176	0.188
0.60	0.129	0.391	0.460	0.017	0.307	0.337
0.70	0.221	0.590	0.665	0.021	0.397	0.468
0.80	0.345	0.786	0.850	0.030	0.605	0.674
0.90	0.505	0.979	0.980	0.045	0.852	0.881



(a) n=20 with 1outlier



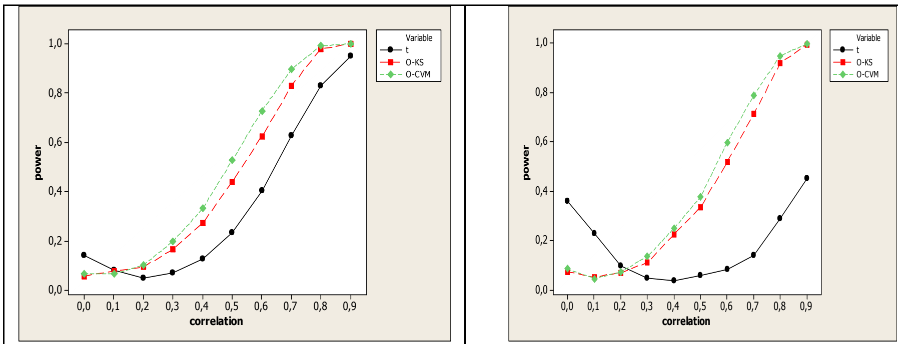
(b) n=20 with 2 outliers

Figure 3. The scatter plots of empirical power estimates of significance tests of correlation

When Table 3 and Figure 4 are examined, the same comments are made as for Table 2 and Figure 3. In addition to these comments, the rates of increase in power for all methods are higher when compared to Table 2 and Figure 3 because of increase in sample sizes.

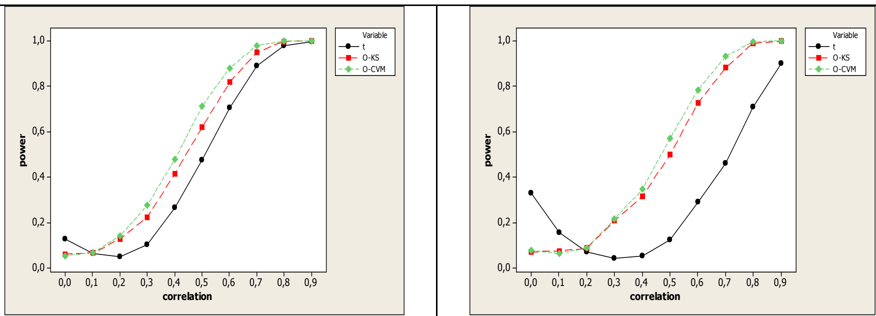
Table 3. The empirical power estimates of significance tests of correlation (1000 repetitions)

n = 30	*mixed with 1 outlier*			*mixed with 2 outliers*		
ρ *values*	*t-test*	*O-KS*	*O-CVM*	*t-test*	*O-KS*	*O-CVM*
0	0.140	0.057	0.066	0.361	0.074	0.085
0.10	0.081	0.077	0.066	0.229	0.052	0.046
0.20	0.047	0.093	0.102	0.097	0.070	0.073
0.30	0.070	0.167	0.198	0.048	0.111	0.136
0.40	0.127	0.272	0.331	0.038	0.224	0.249
0.50	0.234	0.440	0.527	0.058	0.335	0.377
0.60	0.402	0.624	0.728	0.082	0.518	0.598
0.70	0.628	0.830	0.898	0.141	0.712	0.788
0.80	0.829	0.980	0.992	0.289	0.919	0.949
0.90	0.949	1.000	1.000	0.451	0.995	0.996



(a) n=30 with 1outlier (b) n=30 with 1outlier

Figure 4. The scatter plots of empirical power estimates of significance tests of correlation



(a) n=40 with 1outlier (b) n=40 with 2 outliers

Figure 5. The scatter plots of empirical power estimates of significance tests of correlation

Table 4. The empirical power estimates of significance tests of correlation (1000 repetitions)

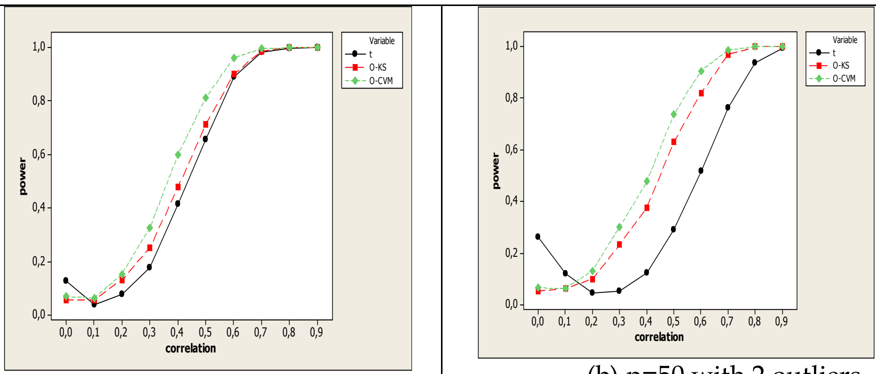
$n = 40$	*mixed with 1 outlier*			*mixed with 2 outliers*		
ρ *values*	*t-test*	*O-KS*	*O-CVM*	*t-test*	*O-KS*	*O-CVM*
0	0.128	0.060	0.053	0.329	0.070	0.076
0.10	0.063	0.066	0.067	0.154	0.074	0.063
0.20	0.049	0.126	0.141	0.070	0.086	0.088
0.30	0.103	0.224	0.276	0.040	0.208	0.216
0.40	0.265	0.415	0.477	0.053	0.314	0.346
0.50	0.474	0.621	0.713	0.124	0.499	0.572
0.60	0.707	0.818	0.880	0.290	0.727	0.784
0.70	0.891	0.950	0.980	0.462	0.882	0.933
0.80	0.978	0.998	0.999	0.710	0.991	0.998
0.90	0.998	1.000	1.000	0.903	1.000	1.000

Again Table 4 and Figure 5 display high Type I error rate and low power for t-test in comparison to the proposed new test approach. It seems that the power for the new test approach reaches acceptable high values for correlation larger than 0.50.

Table 5. The empirical power estimates of significance tests of correlation (1000 repetitions)

$n = 50$	*mixed with 1 outlier*			*mixed with 2 outliers*		
ρ *values*	*t-test*	*O-KS*	*O-CVM*	*t-test*	*O-KS*	*O-CVM*
0	0.125	0.055	0.069	0.261	0.052	0.065
0.10	0.039	0.057	0.063	0.120	0.063	0.063
0.20	0.075	0.131	0.153	0.044	0.099	0.131
0.30	0.176	0.251	0.327	0.051	0.233	0.302
0.40	0.413	0.478	0.599	0.124	0.375	0.477
0.50	0.655	0.712	0.813	0.289	0.630	0.738
0.60	0.891	0.900	0.961	0.518	0.819	0.906
0.70	0.983	0.987	0.996	0.764	0.968	0.986
0.80	0.997	1.000	1.000	0.935	0.999	1.000
0.90	1.000	1.000	1.000	0.995	1.000	1.000

It follows from Table 5 and Figure 6 that despite larger Type I error rate of t-test, its power comes close to the power values of the new test approach in the cases of samples mixed with 1 outlier while superiority of the new test approach in terms of power is more distinct in the case of samples mixed with 2 outliers. The reason for t-test to show such a different performance is the change in the ratio of number of outliers to sample size.



(a) n=50 with 1outlier

(b) n=50 with 2 outliers

Figure 6. The scatter plots of empirical power estimates of significance tests of correlation

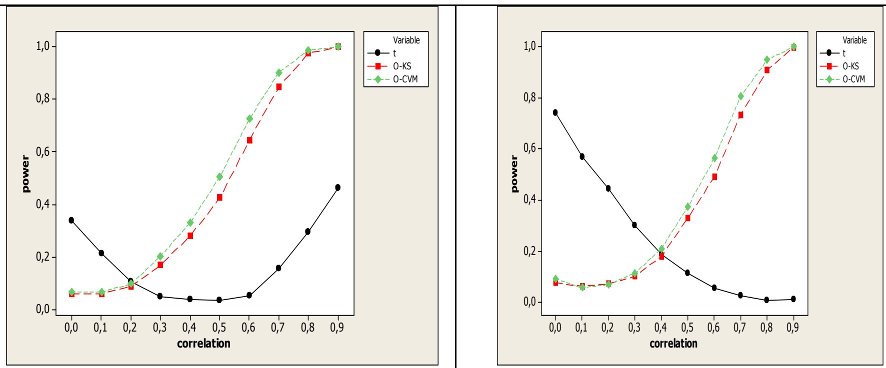
3.1. The Effect of Increasing Magnitude of Outliers

In order to see the effect of increasing magnitude of outliers on the performances of the t-test and the new testing approach, the outliers are generated from the bivariate normal distribution with the means $\mu_X = 3.5$, $\mu_Y = -3.5$, which are larger than the means of the previous outlier distribution 1 units in absolute value. The same patterns of standard deviation and correlation as in data are used in this outlier distribution. The sample size is chosen as 30 to exemplify performances of new testing approach and the t-test in the case of outliers larger in magnitude.

The results are given in Table 6 and displayed in Figure 7. In comparison with the results of Table 3 and Figure 4, performance of the Pearson t-test deteriorates with respect to Type I error probability and power, while the new testing approach maintains its superior and acceptable performance for correlation level above 0.50. Type I error rate of the t-test reaches values at least six times and fourteen times larger than the nominal rate of 0.05 in the cases of samples mixed with 1 and 2 outliers respectively. On the other hand, Type I error of the new approach proposed in this paper exceeds the nominal rate not more than two times.

Table 6. The empirical power estimates of significance tests of correlation (1000 repetitions)

$n=30$	mixed with 1 outlier			mixed with 2 outliers		
ρ *values*	*t-test*	*O-KS*	*O-CVM*	*t-test*	*O-KS*	*O-CVM*
0	0.336	0.060	0.065	0.739	0.076	0.090
0.10	0.211	0.059	0.066	0.566	0.062	0.057
0.20	0.105	0.089	0.097	0.443	0.072	0.069
0.30	0.047	0.169	0.203	0.302	0.101	0.115
0.40	0.036	0.279	0.330	0.187	0.18	0.209
0.50	0.035	0.425	0.504	0.114	0.329	0.374
0.60	0.053	0.644	0.726	0.053	0.491	0.564
0.70	0.155	0.847	0.902	0.026	0.734	0.806
0.80	0.293	0.975	0.987	0.008	0.907	0.947
0.90	0.461	1.000	1.000	0.009	0.998	0.999



(a) n=30 with 1outlier

(b) n=30 with 2 outliers

Figure 7. The scatter plots of empirical power estimates of significance tests of correlation with outlier/s larger in magnitude (n=30)

4. RESULTS AND DISCUSSIONS

Under the conditions considered in the simulation study, first, it is worthy of notice that Type I error rate of the Pearson t-test for correlation is unacceptably high: it tends to be quite large than the specified α level and the Type I error rate of the new testing approach. Further, it appears that Type I error rate of the t-test inflates as the outliers increase in magnitude or in number. Second, the power of t-test tends to be lower than that of the new testing approach. As the outliers increase further in magnitude, the power of the t-test decreases compared to that of the new testing approach. On the other hand, the performance of the proposed testing approach does not seem to be affected by the change in the magnitude of outliers.

The superior performance of the new testing approach over the Pearson t-test in terms of Type I error rate and power decreases as the outliers increase in number: the nice power properties as well as Type I error rate of the new testing approach deteriorate in such cases. Thus, it seems that the proposed testing approach in this paper is a plausible alternative to the t-test as long as outliers constitute at most approximately 5% or 6% of the sample.

The possible reason for lower performance of the t-test in the case of outliers is its dependence on testing whether the observed value of the single statistics, namely sample estimate r of Pearson product moment correlation, is a usual observation from the distribution of the statistics under the null hypothesis. In contrast, the new testing approach proposed in this paper is based on testing whether the whole observed sample comes from the relevant distribution under the null hypothesis. Thus, the new testing approach can alleviate negative effect of outliers unless they are large in number.

To sum up, the new testing approach proposed outperforms the Pearson t-test for correlation in terms of Type I error rates and power for the conditions considered in this study.

SUGGESTIONS

In view of findings in this study, the new testing approach proposed in this paper should be considered as an alternative to the usual Pearson t-test for correlation under the conditions considered in this study. The reason for this is that the t-test breaks down in terms of Type I error and power, which can be regarded as the measure of the ability of making right decisions for a test procedure.

REFERENCES

Bain, L. J. & Engelhardt, M. 1992. Introduction to Probability and Mathematical Statistics (2nd ed.), Duxbury Press, Belmont, California, 644p.

Bishara, A. J., & 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.

Conover, W.J. 1999. Practical Nonparametric Statistics (3rd ed.), John Wiley & Sons, New York, 592p.

El-Fallah, M., & El-Salam, A. 2006. A robust measure for the correlation coefficient. J. King Saud Univ., 19, Admin. Sci. 1, 41-50.

Evandt O., Coleman, S., Ramalhoto M.F. & Lottum, C.V. Y.J. 2004. A little known Robust Estimator of the Correlation Coefficient and Its Use in a Robust Graphical Test for Bivariate Normality with Applications in Aluminum Industry, Mathematical Programming, Qual.Reliab.Engng. Int. 20, 433-456.

Huber, P. J. 2004. Robust Statistics, John Wiley & Sons, USA, 308p.

King, J. E. 2003. What have we learned from 100 years of robustness studies on r? Paper presented at the annual meeting of the Southwest Educational Research Association, San Antonio, Texas, US. February.

Pernet CR, Wilcox, R. and Rousselet, G.A. 2013. Robust correlation analyses: false positive and power validation using a new open source Matlab toolbox. Front. Psychology 3:606. doi: 10.3389/fpsyg.2012.00606

Scott, W.F. 2000. Tables of Cramer-Von Mises Distributions, Commun.Statist.-Theory Meth. 29(1), 227-235.

Stephens, M.A. 1974. EDF Statistics for Goodness-of-Fit and Some Comparisons, J.Amer.Statist.Assoc. 69, 730-737.

Stephens, M.A. 1977. Goodness-of-Fit for the Extreme-Value Distribution, Biometrika 64, 583-588.

Wackerly D.D., Mendenhall, W. & Scheaffer, R.L. 2007. Mathematical Statistics with Applications (7th ed.), Australia, Brooks/Cole,896p.

Wilcox, R. R. 2012. Introduction to robust estimation and hypothesis testing(3rd ed.), Academic Press, USA, 690p.

****