COMPARING PEARSON CORRELATIONS:
DEALING WITH HETEROSCEDASTICITY AND
NON-NORMALITY

Rand R. Wilcox
Dept of Psychology
University of Southern California

August 21, 2009
ABSTRACT

Methods for computing a confidence interval for the difference between two Pearson correlations are compared when dealing with non-normality and heteroscedasticity. Variations of a method derived by Zou performed relatively well in simulations for the dependent case. For the independent case, the Wilcox-Muska method performed best.

Keywords: Measures of association, level robust methods, non-normality, heteroscedasticity.

AMS Classifications: 62G09, 62F03, 6205

1 Introduction

Consider two Pearson correlations, $r_1$ and $r_2$, which may or may not be independent, and denote the corresponding population values by $\rho_1$ and $\rho_2$. Numerous papers have taken up the problem of testing

$$H_0 : \rho_1 = \rho_2,$$

the hypothesis that the two correlations are equal (e.g., Hittner, May & Silver, 2003; Hotelling, 1940; Olkin, 1967; Dunn & Clark, 1969; Meng, Rosenthal & Rubin, 1992; Steiger, 1980; Wilcox & Muska, 2002; Wilcox & Tian, 2008; Williams, 1959; Zou, 2007.) Among extant techniques, most are known to be unsatisfactory, in terms of controlling Type I error probabilities under non-normality, or heteroscedasticity, or both (e.g., Hittner, May & Silver, 2003; Wilcox, 2005). In particular, all methods based on Fisher’s z are suspect due to both theoretical results (Duncan & Layard, 1973) and simulation studies (e.g., Hittner, May & Silver, 2003; Berry and Mielke, 2000).

Recently, Zou (2007) derived a method aimed at the important problem of computing a confidence interval for $\rho_1 - \rho_2$. Zou went on to suggest that a particular variation be used, which is based in part on Fisher’s r-to-z transformation. To be a bit more precise, Fisher’s $z$ is used to compute confidence intervals for the individual correlations, which in turn can be used to compute a confidence interval for $\rho_1 - \rho_2$. (Details are reviewed later the paper.) Although there are known concerns regarding Fisher’s $z$, the basic structure of Zou’s method does not require that Fisher’s $z$ be used; a more level robust approach can be used to get confidence intervals for the individual correlations, which provides an alternative approach to computing a confidence interval for $\rho_1 - \rho_2$. And of course, this means that in particular, (1) can be tested. Currently, it seems that three alternative methods for computing a confidence
Comparing Correlations

Before continuing, it is noted when testing hypotheses about Pearson’s correlation with some bootstrap method, as opposed to computing a confidence interval, the issue of resampling under specific hypotheses, with the goal of controlling the level of the test and estimating power, becomes relevant. Martin (2007) provides a recent discussion of this issue and extensions of his results might practical value for the problem at hand. Here, however, this possibility is left to future investigations.

Consider four random variables: $Y_1, Y_2, X_1$ and $X_2$. When comparing dependent correlations, two situations are of interest. The first is the so-called overlapping case where $\rho_j$ is the correlation between $X_j$ and $Y_1$, $j = 1, 2$. In the non-overlapping case, $\rho_j$ is the correlation between $X_j$ and $Y_j$, where $Y_1$ and $Y_2$, as well as $X_1$ and $X_2$, are possibly dependent. Wilcox and Tian (2008) suggested a method for dealing with the overlapping case. Their method offers some protection against non-normality and it appears to improve upon all of the methods compared by Hittner et al. (2003). However, results reported here indicate that a modification of Zou’s method provides reasonably good control over the probability of a Type I error for a broader range of situations. The modification again replaces Fisher’s $z$ with an alternative method for computing a confidence interval for the individual correlations. But unlike the independent case, the modification stems from some recent results on confidence intervals for the slope of a least squares regression line. (This is method HC4 described later in the paper.)

2 Computing a Confidence Interval for a Single Correlation

Momentarily consider two random variables only, $X$ and $Y$, having correlation $\rho$. This section reviews some key results regarding confidence intervals for $\rho$ that will be needed in this paper.

A Basic Percentile Bootstrap Method

Based on a random sample of $n$ pairs of points, $(X_1, Y_1), \ldots, (X_n, Y_n)$, if the goal is to compute a $1 - \alpha$ confidence interval for $\rho$, the basic percentile bootstrap method (e.g., Efron & Tibshirani, 1993; Liu & Singh, 1997) proceeds as follows:

1. Generate a bootstrap sample by resampling with replacement, $n$ pairs of points from $(X_1, Y_1), \ldots, (X_n, Y_n)$. 

interval for a single correlation deserve serious consideration, which are reviewed in the next section of this paper.
2. Compute Pearson’s correlation based on this bootstrap sample yielding $r^*$.  

3. Repeat steps 1 and 2 $B$ times yielding $r^*_1, \ldots, r^*_B$.  

4. Put the values $r^*_1, \ldots, r^*_B$ is ascending and label the results $r^*_1 \leq \ldots \leq r^*_B$.  

5. Let $\ell = \alpha B/2$, rounded to the nearest integer and $u = B - \ell$. Then the $1 - \alpha$ confidence interval for $\rho$ is  

$$(r_{(\ell+1)}, r_{(u)}).$$  

(2)

**A Modified Percentile Bootstrap Method**

A concern with the basic percentile bootstrap method is that when computing a .95 confidence interval, and when the sample size is small, the actual probability coverage can be substantially smaller than .95. To correct this problem, Wilcox and Muska (2001) first determined an adjustment under normality and homoscedasticity. (So the spirit of the method has obvious similarities to Gosset’s derivation of Student’s T.) Their adjusted .95 confidence interval for $\rho$, using $B = 599$ bootstrap samples, is  

$$(r^*_a, r^*_c)$$  

(3)

where for $n < 40$, $a = 7$ and $c = 593$; for $40 \leq n < 80$, $a = 8$ and $c = 592$; for $80 \leq n < 180$, $a = 11$ and $c = 588$; for $180 \leq n < 250$, $a = 14$ and $c = 585$; while for $n \geq 250$, $a = 15$ and $c = 584$. For $n \geq 250$, a standard percentile bootstrap method is used.

The choice of $B = 599$, rather than $B = 600$ stems from results indicating that when a test statistics is pivotal, $B$ should be chosen so that $\alpha(B + 1)$ is an integer. (See, for example, Racine and MacKinnon, 2007a.) (For larger values of $B$, simply adjust $a$ and $c$ so that $a/B$, for example, is approximately equal to the ratio obtained when $B = 599$.) This will be called modification M1. Simulations reported by Wilcox and Muska (2001) indicate that method M1 performs well when sampling from non-normal distributions or when there is heteroscedasticity, and this is the main motivation for considering it here. (An alternative approach that also widens the confidence intervals was suggested by Efron, 1982, pp. 84-86, which was studied by Strube, 1988.) A criticism of using not using a larger value for $B$ is that this might lead to some loss of power when the goal is to test (1). Racine and MacKinnon (2007a) discuss this issue at length and proposed a method for reducing this problem. (Also see Jökel, 1986). Davidson and MacKinnon (2000) proposed a pretest procedure for choosing $B$.

Now, a natural variation of Zou’s method, when trying to deal with non-normality and heteroscedasticity, is to use M1 to compute a confidence interval for the individual correlations rather than Fisher’s z. But simulations done here indicate that this approach is not
quite satisfactory when \( n \) is large. What is better is to use method M1, except when \( n \geq 180 \), in which case use \( a = 14 \) and \( c = 585 \) in equation (3). This will be called method M2.

**The HC4 Method**

Momentarily consider the regression model \( Y = \beta_0 + \beta_1 X + \lambda(x)e \), where \( X \) and \( e \) are independent and \( \lambda \) is some function used to model heteroscedasticity. (So \( \lambda(X) \equiv 1 \) corresponds to the usual homoscedastic model.) Let \( b_0 \) and \( b_1 \) be the usual least squares estimates of \( \beta_0 \) and \( \beta_1 \), respectively. A fundamental concern associated with the usual confidence intervals for \( \beta_0 \) and \( \beta_1 \) is that, under general conditions, invalid estimates of the standard errors of \( b_0 \) and \( b_1 \) are used when there is heteroscedasticity. Recently, however, there has been substantial progress in terms of computing confidence intervals when there is heteroscedasticity (e.g., Cribari-Neto, 2004; Godfrey, 2006; Liu, 1988; Davidson & Flachaire, 2000; Davidson & MacKinnon, 2004) using estimates of the standard errors that remain valid under heteroscedasticity. Several alternative estimates of the standard error have been proposed. An early and important alternative is the HC0 estimator derived by White (1980). More recently, MacKinnon and White (1985) introduced several alternatives to HC0, typically labeled HC1, HC2 and HC3. This was followed by the HC4 estimator derived by Cribari-Neto (2004). Also see Cribari-Neto, Souza and Vasconcellos, (2007). Here the focus is on the HC4 estimate of the squared standard errors, which is motivated in part by results in Godfrey (2006). This is not to suggest, however, that the HC4 estimator dominates all other heteroscedastic-consistent estimators for the problem at hand. Presumably this is not the case, but this issue is left for future investigations.

Let \( X \) be the \( n \)-by-2 matrix with ones in the first column and the predictor values, \( X \), in the second. Let

\[
V = (X'X)^{-1}X'diag\left[\frac{r_i^2}{(1-h_{ii})^b_i}\right]X(X'X)^{-1},
\]

where \( r_1, \ldots, r_n \) are the usual residuals,

\[
\delta_i = \min \left\{ 4, \frac{nh_{ii}}{\sum_{i=1}^{n} h_{ii}} \right\}
\]

\[
h_{ii} = x_i(X'X)^{-1}x_i'.
\]

Moreover,

\[
\frac{b_1 - \beta_1}{\sqrt{V_{22}}}.
\]

is asymptotically normal (e.g., Cribari-Neto, 2004), and this suggests a simple way of getting a confidence interval for \( \rho \): standardize the \( X \) values, in which case \( b_1 \) is the usual estimate of \( \rho \), \( r \), and use

\[
b_1 \pm c\sqrt{V_{22}},
\]
where \( c \) is the \( 1 - \alpha/2 \) of a standard normal distribution. And as will become evident, this provides yet another variation of Zou’s method for testing (1).

### 3 The Wilcox-Muska Confidence Interval for \( \rho_1 - \rho_2 \), \( r_1 \) and \( r_2 \) Independent

For the special case where \( r_1 \) and \( r_2 \) are independent, Wilcox and Muska (2002) found that a simple extension of the modified percentile bootstrap for computing a confidence interval for \( \rho \) to be relatively effective when computing a .95 confidence interval for \( \rho_1 - \rho_2 \). Like method M1, the derivation of the method was based on the simple strategy of making adjustments to the confidence, stemming from the basic percentile bootstrap method, when sampling from a bivariate normal distribution. In particular, take a bootstrap sample from each group, compute Pearson’s correlation and label the results \( r^*_1 \) and \( r^*_2 \), and let \( D^* = r^*_1 - r^*_2 \). Repeat this process \( B \) times yielding \( D^*_1, \ldots, D^*_B \) and put these values in ascending order yielding \( D^*_1 \leq \cdots \leq D^*_B \). Letting \( n_1 \) and \( n_2 \) denote the sample sizes, let \( N = n_1 + n_2 \). The .95 confidence interval for \( \rho_1 - \rho_2 \) is

\[
(D^*_{(\ell)}), \ D^*_{(u)}),
\]

where for \( \ell = 7 \) and \( u = 593 \) if \( N < 40 \); \( \ell = 8 \) and \( u = 592 \) if \( 40 \leq N < 80 \); \( \ell = 11 \) and \( u = 588 \) if \( 80 \leq N < 180 \); \( \ell = 14 \) and \( u = 585 \) if \( 180 \leq N < 250 \); \( \ell = 15 \) and \( u = 584 \) if \( N \geq 250 \). As is evident, this method is very similar to method M1 and it reduces to the standard percentile bootstrap method when \( N \geq 250 \). (A similar modified bootstrap method has been considered when computing a .99 confidence interval, but now probability coverage can be somewhat unsatisfactory when dealing non-normality and heteroscedasticity, at least with small to moderate sample sizes.)

### 4 Zou’s Method

For convenience, the details of Zou’s method are summarized here. (For related results motivating this method, see Donner & Zou, 2008.)

**Independent Correlations**

For the case where \( r_1 \) and \( r_2 \) are independent, let \((l_1, u_1)\) and \((l_2, u_2)\) be \( 1 - \alpha \) confidence intervals for \( \rho_1 \) and \( \rho_2 \), respectively. Then a \( 1 - \alpha \) confidence interval for \( \rho_1 - \rho_2 \) is \((L, U)\),
and $(j = 1$.

Terminating the confidence intervals $(l_1, u_1)$ and $(l_2, u_2)$ be 1 – $\alpha$ confidence intervals for $\rho_{12}$ and $\rho_{13}$, respectively. Now

\[
L = r_{12} - r_{13} - \sqrt{(r_{12} - l_1)^2 + (u_2 - r_{13})^2 - 2\rho_{12}r_{13}(r_{12} - l_1)(u_2 - r_{13})},
\]

and

\[
U = r_{12} - r_{23} + \sqrt{(u_1 - r_{12})^2 + (r_{23} - l_2)^2 - 2\rho_{12}r_{13}(r_{12} - l_1)(r_{23} - l_2)}.
\]

where

\[
\texttt{corr}(r_{12}, r_{13}) = \frac{(r_{23} - \cdot 5r_{12}r_{23})(1 - r_{12}^2 - r_{13}^2 - r_{23}^2) + r_{24}^2}{(1 - r_{12}^2)(1 - r_{13}^2)}.
\]

Nonoverlapping Correlations

Again following Zou’s notation, for the overlapping case, $\rho_{jk}$ is the correlation between $X_j$ and $X_k$, $j = 1, 2, 3, 4$; $k = 1, 2, 3, 4$ and the goal is to compute a confidence interval for $\rho_{12} - \rho_{34}$. Now $(l_1, u_1)$ and $(l_2, u_2)$ are 1 – $\alpha$ confidence intervals for $\rho_{12}$ and $\rho_{34}$, respectively. And the 1 – $\alpha$ confidence interval for $\rho_{12} - \rho_{34}$ is $(L, U)$, where

\[
L = r_{12} - r_{34} - \sqrt{(r_{12} - l_1)^2 + (u_2 - r_{34})^2 - 2\rho_{12}r_{13}(r_{12} - l_1)(u_2 - r_{34})},
\]

\[
U = r_{12} - r_{34} + \sqrt{(u_1 - r_{12})^2 + (r_{34} - l_2)^2 - 2\rho_{12}r_{13}(r_{12} - l_1)(r_{34} - l_2)},
\]

where

\[
\texttt{corr}(r_{12}, r_{13}) = \frac{T1 - T2}{T3},
\]

\[
T1 = .5r_{12}r_{34}(r_{13}^2 - r_{14}^2 + r_{23}^2 + r_{24}^2) + r_{13}r_{24} + r_{14}r_{23},
\]

\[
T2 = r_{12}r_{13}r_{14} + r_{12}r_{23}r_{24} + r_{13}r_{23}r_{34} + r_{14}r_{24}r_{34}
\]

and

\[
T3 = (1 - r_{12}^2)(1 - r_{34}^2).
\]

As previously indicated, for all three of the methods just described, Zou suggests determining the confidence intervals $(l_1, u_1)$ and $(l_2, u_2)$ via Fisher’s z. But as will be seen
this approach can be highly unsatisfactory, even under normality. The issue is the extent replacing Fisher’s z with methods M2 or HC4 gives improved results.

For completeness, it is noted that yet another method was considered when dealing with the overlapping case. Bootstrap samples are generated as done by the Wilcoxon-Muska method, but a p-value was estimated based on a kernel density estimate of the bootstrap sampling distribution, a strategy motivated by results in Racine and MacKinnon (2007b). Several other modifications were considered as well, the details of which are given in Wilcox (2008). A positive feature is that it was found to improve on the Wilcoxon-Tian method, but here it was found that the HC4 variation of Zou’s method provided more accurate probability coverage, so the rather involved computational details are omitted. Also, Beasley, DeShea, Toothaker, Mendoza, Bard and Rodgers (2007) suggest yet another bootstrap method for computing a confidence interval for $\rho$, it was considered here, but it performed poorly under heteroscedasticity and so it was abandoned.

5 Design of the Simulation Study

Four types of distributions are considered: normal, symmetric and heavy-tailed (roughly meaning that outliers tend to be common), asymmetric and relatively light-tailed, and asymmetric and relatively heavy-tailed. More specifically, g-and-h distributions (Hoaglin, (1985) are used, which arise as follows. Let $Z$ be a random variable having a standard normal distribution. Then

$$W = \begin{cases} \frac{\exp(gZ) - 1}{g} \exp(hZ^2/2), & \text{if } g > 0 \\ Z \exp(hZ^2/2), & \text{if } g = 0 \end{cases}$$

has a g-and-h distribution, where $g$ and $h$ are parameters that determine the first four moments. The four distributions use here are the standard normal ($g = h = 0$), a symmetric heavy-tailed distribution ($h = .2, g = 0$), an asymmetric distribution with relatively light tails($h = 0, g = .2$), and an asymmetric distribution with heavy tails ($g = h = .2$). Table 1 summarizes the skewness ($\gamma_1$) and kurtosis ($\gamma_2$) of these distributions.

Two sets of simulations were run. The first is where $X$ has a normal distribution and the error term has one of the g-and-h distributions just described. For the second set, $X$ has the same distribution as the error term.

For independent case, observations were generated for $X_1$ and $X_2$, as well as $e_1$ and $e_2$, from one of the four g-and-h distributions previously described, with $X_1$ and $X_2$ independent of $e_1$ and $e_2$, yielding $Y_1 = \beta X_1 + \lambda(X_1)e_1$ and $Y_2 = \beta X_2 + \lambda(X_2)e_2$. The choices for $\lambda(X)$ were taken to be $\lambda(X) \equiv 1$, $\lambda(X) = |X| + 1$ (so the conditional variance of $Y$, given $X$,}
Comparing Correlations

is smallest when \( X \) is close to its mean), and \( \lambda(X) = 1/(|X| + 1) \) (in which case the the conditional variance of \( Y \), given \( X \), is largest when \( X \) is close to its mean. For convenience these three choices for \( \lambda \) will be called variance patterns (VP) 1, 2 and 3, respectively. For the dependent cases, the correlation between \( X_1 \) and \( X_2 \) was taken to be 0 or .5. For the dependent overlapping case, \( Y = \beta X_1 + \beta X_2 + \lambda(X_1, X_2)e \), where now VP 2 and 3 refer to \( \lambda(X_1, X_2) = (|X_1 + X_2| + 1) \) and \( \lambda(X_1, X_2) = 1/(|X_1 + X_2|) + 1 \), respectively. For the overlapping case, \( Y_j = \beta X_j + \lambda(X_j)e_j \) \((j = 1, 2)\), with VP 2 and VP 3 defined as in the independent case. The two choices for \( \beta \) were 0 and 1.

Values for \( X_1 \) and \( X_2 \) were generated with the R function rmul, which can be used to generate multivariate data having specified marginal distributions with specified Pearson correlations. This function has an argument that allows the user to specify the marginal distributions, and here the function ghdist was used, which generates data from a g-and-h distribution. Both rmul and ghdist are stored in a library of functions that can be downloaded from the author’s web page listed in the final section of this paper.

6 Simulation Results

Independent Case

First consider the case where \( r_1 \) and \( r_2 \) are independent. Simulations based on 5000 replications were used to estimate the actual Type I error probability when testing at the \( \alpha = .05 \) level. So if the actual Type I error probability is .05, the standard error of the estimate is only .003.

Table 2 shows the estimated Type I error probabilities when \( n_1 = n_2 = 30 \). Additional simulations were run with \( n_1 = 20 \) and \( n_2 = 40 \), the pattern of results are similar to those in Table 2, so they are not reported. When \( \beta = 0 \), there is little separating the various methods. But when \( \beta = 1 \), certain methods break down, even under normality, and so for brevity, attention is focused on \( \beta = 1 \). Results using Zou’s method with Fisher’s z are labeled ZF, Zou’s method with the M2 bootstrap method are labeled ZM2, and results using the Wilcoxon-Muska method are labeled WM. It is evident that under heteroscedasticity, method ZF can be highly unsatisfactory. Generally methods ZM2 and WM perform reasonable well. Typically, ZM2 results in estimated Type I errors less than those obtained with WM, suggesting WM might have a bit more power. Note that for WM, the estimated Type I error probabilities range between .020 and .053. Using method HC4, the estimated Type I error probability never exceeded the nominal level by very much. But a concern is that the estimate can be substantially smaller than the nominal level in some cases, generally
Comparing Correlations

ZM2 and WM give better results, so results on HC4 are omitted. Notice that even under normality and homoscedasticity, method ZF has an estimated Type I error probability of .005. This is in contrast to $\beta = 0$ where the estimate is .046. Also note that method ZF has estimates as high as .2 when both $X$ and $e$ have the same g-and-h distribution. All indications are that in terms of Type I errors, ZM2 and WM are the only methods that perform reasonably well over a broad range of situations.

Overlapping Case

Next, consider the case where $r_1$ and $r_2$ are dependent and first focus on the overlapping case. To begin, consider the case where $\beta = 1$ and the correlation between $X_1$ and $X_2$ is $\rho(X_1, X_2) = 0$. Table 3 shows the estimated Type I errors when testing at the .05 level using ZF and HC4. Execution time using the alternative bootstrap methods proved to be extremely high. Simulations based on 1000 replications suggest using HC4, and so here results using HC4 are reported that are based on 5000 replications. Of course, a larger number or replications might reveal that one of the bootstrap methods has practical value, but this is not pursued here. As can be seen, method HC4 performs reasonably well over all of the conditions considered, in terms of avoiding Type I errors well above the nominal level. In fairness, the same is generally true of ZF, but in one situation the estimate is .175.

Table 4 shows the results where now $\beta = 0$. Note that method ZF now performs poorly in some situations where it performed well when $\beta = 1$. Again method HC4 avoids Type I errors well above the nominal level. Its main difficulty is that in some cases the estimate drops below .02. Table 5 shows the results when again $\beta = 0$, only now $\rho(X_1, X_2) = .5$. Again, HC4 avoids Type I errors well above the nominal level, the highest estimate being .059. Finally, Table 6 shows the results when $\beta = 1$ and $\rho(X_1, X_2) = .5$. Again, HC4 avoids Type I errors well above the nominal level, but often the estimate is well below the nominal level. So HC4 is recommended for general use, but clearly an improved method would be desirable.

The Non-Overlapping Case

As for the non-overlapping case, HC4 generally performs well, but now there are various situations where the estimated Type I error probability can be well above or below the nominal level. All indications are that ZM2 is much more satisfactory, so now the focus is on ZM2. (Simulation results using HC4 are available upon request.) The estimated Type I error probabilities are reported in Tables 7-10 for the same situations considered in Tables
3-6. Again there are situations where the estimated Type I error probability drops below .025, but in general the method appears to perform reasonably well, the highest estimate being .053.

7 Concluding Remarks

In summary, when comparing independent correlations, the results reported here suggest using method WM. For the overlapping case, use HC4, and for the non-overlapping case use ZM2. Software for applying these methods is stored in a library of R functions that can be downloaded from www-rcf.usc.edu/~rwilcox/. (Download the file Rallfun-v9. The functions that apply methods WM, HC4 and ZM2 are twopcor, TWOpov and TWOpNOV, respectively.

To put this paper in a broader perspective, it is noted that two general goals need additional attention. The first is comparing robust dependent correlations, particularly when using a robust regression estimator. It is well known that Pearson’s correlation is not robust (e.g., Wilcox, 2005). In practical terms, it can be highly misleading in terms of assessing the strength of an association, and when testing hypotheses, power can be relatively low versus using some robust analog. Many robust correlations have been proposed as well as many robust regression estimators. Methods for comparing independent robust measures of association have been derived, but little is known about how best to address the dependent case. Yet another concern is curvature. It is known that reliance on standard parametric models for dealing with curvature are not always satisfactory. So more flexible approaches for dealing with curvature are needed when comparing the strength of the association of two predictors. Methods for dealing with these issues are under investigation.

Finally, it is reiterated that there are variations of the methods described here that might have practical value when testing (1). Of interest, particularly when assessing power, are results reported by Martin (2007).
References


Comparing Correlations

formula for significance levels of $r_{13}$ vs. $r_{23}$ compared with Hotelling’s method. American Educational Research Journal, 7, 189–195.


Table 1: Some properties of the g-and-h distribution

<table>
<thead>
<tr>
<th>g</th>
<th>h</th>
<th>γ_1</th>
<th>γ_2</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0</td>
<td>0.0</td>
<td>0.00</td>
<td>3.00</td>
</tr>
<tr>
<td>0.0</td>
<td>0.2</td>
<td>0.00</td>
<td>21.46</td>
</tr>
<tr>
<td>0.2</td>
<td>0.0</td>
<td>0.61</td>
<td>3.68</td>
</tr>
<tr>
<td>0.2</td>
<td>0.2</td>
<td>2.81</td>
<td>155.98</td>
</tr>
</tbody>
</table>


Table 2: Estimated Type I error probabilities when comparing independent correlations, $\beta = 1, n_1 = n_2 = 30, \alpha = .05$.

<table>
<thead>
<tr>
<th>g</th>
<th>h</th>
<th>VP</th>
<th>ZF</th>
<th>ZM2</th>
<th>WM</th>
<th>ZF</th>
<th>ZM2</th>
<th>WM</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0</td>
<td>0.0</td>
<td>1</td>
<td>.005</td>
<td>.023</td>
<td>.036</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>.079</td>
<td>.038</td>
<td>.044</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>.000</td>
<td>.020</td>
<td>.023</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>.035</td>
<td>.032</td>
<td>.039</td>
<td>.025</td>
<td>.032</td>
<td>.039</td>
</tr>
<tr>
<td>2</td>
<td>.089</td>
<td>.032</td>
<td>.040</td>
<td>.167</td>
<td>.046</td>
<td>.053</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>.014</td>
<td>.030</td>
<td>.035</td>
<td>.003</td>
<td>.032</td>
<td>.044</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>.006</td>
<td>.029</td>
<td>.036</td>
<td>.003</td>
<td>.029</td>
<td>.038</td>
</tr>
<tr>
<td>2</td>
<td>.088</td>
<td>.033</td>
<td>.041</td>
<td>.078</td>
<td>.031</td>
<td>.048</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>.000</td>
<td>.021</td>
<td>.026</td>
<td>.000</td>
<td>.024</td>
<td>.039</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.0</td>
<td>0.2</td>
<td>1</td>
<td>.041</td>
<td>.032</td>
<td>.039</td>
<td>.043</td>
<td>.042</td>
<td>.049</td>
</tr>
<tr>
<td>2</td>
<td>.110</td>
<td>.032</td>
<td>.201</td>
<td>.034</td>
<td>.044</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>.013</td>
<td>.031</td>
<td>.040</td>
<td>.016</td>
<td>.039</td>
<td>.051</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

ZF: Zou’s method with Fisher’s z, ZM2: Zou’s method with M2, WM: Wilcox-Muska method

Table 3: Estimated Type I error probabilities, dependent overlapping case, $\beta = 1, \rho(X_1, X_2) = 0, n = 30$

<table>
<thead>
<tr>
<th>g</th>
<th>h</th>
<th>VP</th>
<th>ZF</th>
<th>HC4</th>
<th>ZF</th>
<th>HC4</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0</td>
<td>0.0</td>
<td>1</td>
<td>.021</td>
<td>.017</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>.018</td>
<td>.024</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>.012</td>
<td>.013</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>.028</td>
<td>.016</td>
<td>.028</td>
<td>.018</td>
</tr>
<tr>
<td>2</td>
<td>.046</td>
<td>.015</td>
<td>.058</td>
<td>.018</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>.016</td>
<td>.014</td>
<td>.020</td>
<td>.025</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>.021</td>
<td>.015</td>
<td>.054</td>
<td>.041</td>
</tr>
<tr>
<td>2</td>
<td>.040</td>
<td>.015</td>
<td>.072</td>
<td>.021</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>.013</td>
<td>.012</td>
<td>.046</td>
<td>.059</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>.029</td>
<td>.016</td>
<td>.061</td>
<td>.021</td>
</tr>
<tr>
<td>2</td>
<td>.044</td>
<td>.014</td>
<td>.175</td>
<td>.016</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>.016</td>
<td>.014</td>
<td>.019</td>
<td>.020</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

ZF: Zou’s method with Fisher’s z, HC4: Zou’s method with HC4
Table 4: Estimated Type I error probabilities, dependent overlapping case, $\beta = 0$, $\rho(X_1, X_2) = 0, n = 30$

<table>
<thead>
<tr>
<th>g</th>
<th>h</th>
<th>VP</th>
<th>ZF</th>
<th>HC4</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0</td>
<td>0.0</td>
<td>1</td>
<td>.052</td>
<td>.042</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>.040</td>
<td>.015</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>.012</td>
<td>.011</td>
</tr>
<tr>
<td>0.0</td>
<td>0.2</td>
<td>1</td>
<td>.084</td>
<td>.047</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>.147</td>
<td>.019</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>.079</td>
<td>.067</td>
</tr>
<tr>
<td>0.2</td>
<td>0.0</td>
<td>1</td>
<td>.027</td>
<td>.019</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>.057</td>
<td>.018</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>.016</td>
<td>.016</td>
</tr>
<tr>
<td>0.2</td>
<td>0.2</td>
<td>1</td>
<td>.102</td>
<td>.054</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>.173</td>
<td>.018</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>.097</td>
<td>.076</td>
</tr>
</tbody>
</table>

ZF: Zou’s method with Fisher’s z  
ZM2: Zou’s method with M2  
WT: Wilcox-Tian method  
HC4: Zou’s method with HC4
Table 5: Estimated Type I error probabilities, dependent overlapping case, $\beta = 0$, $\rho(X_1, X_2) = 0.5$, $n = 30$

<table>
<thead>
<tr>
<th>g</th>
<th>h</th>
<th>VP</th>
<th>ZF</th>
<th>HC4</th>
<th>ZF</th>
<th>HC4</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0</td>
<td>0.0</td>
<td>1</td>
<td>.049</td>
<td>.040</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>.042</td>
<td>.008</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>.042</td>
<td>.010</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.0</td>
<td>0.2</td>
<td>1</td>
<td>.048</td>
<td>.025</td>
<td>.049</td>
<td>.027</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>.019</td>
<td>.054</td>
<td>.126</td>
<td>.008</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>.137</td>
<td>.007</td>
<td>.020</td>
<td>.050</td>
<td></td>
</tr>
<tr>
<td>0.2</td>
<td>0.0</td>
<td>1</td>
<td>.051</td>
<td>.038</td>
<td>.049</td>
<td>.038</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>.043</td>
<td>.011</td>
<td>.063</td>
<td>.010</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>.063</td>
<td>.010</td>
<td>.046</td>
<td>.059</td>
<td></td>
</tr>
<tr>
<td>0.2</td>
<td>0.2</td>
<td>1</td>
<td>.050</td>
<td>.023</td>
<td>.055</td>
<td>.023</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>.018</td>
<td>.024</td>
<td>.175</td>
<td>.017</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>.157</td>
<td>.017</td>
<td>.017</td>
<td>.020</td>
<td></td>
</tr>
</tbody>
</table>

XN: $X$ has bivariate normal distribution
XGH: $X$ has a bivariate g-and-h distribution
ZF: Zou’s method with Fisher’s $z$
HC4: Zou’s method with HC4

Table 6: Estimated Type I error probabilities, dependent overlapping case, $\beta = 1$, $\rho(X_1, X_2) = 0.5$, $n = 30$

<table>
<thead>
<tr>
<th>g</th>
<th>h</th>
<th>VP</th>
<th>ZF</th>
<th>HC4</th>
<th>ZF</th>
<th>HC4</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0</td>
<td>0.0</td>
<td>1</td>
<td>.002</td>
<td>.004</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>.002</td>
<td>.001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>.020</td>
<td>.007</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.0</td>
<td>0.2</td>
<td>1</td>
<td>.002</td>
<td>.003</td>
<td>.004</td>
<td>.003</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>.001</td>
<td>.001</td>
<td>.085</td>
<td>.006</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>.070</td>
<td>.005</td>
<td>.001</td>
<td>.001</td>
<td></td>
</tr>
<tr>
<td>0.2</td>
<td>0.0</td>
<td>1</td>
<td>.002</td>
<td>.005</td>
<td>.002</td>
<td>.005</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>.000</td>
<td>.004</td>
<td>.033</td>
<td>.008</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>.033</td>
<td>.007</td>
<td>.000</td>
<td>.001</td>
<td></td>
</tr>
<tr>
<td>0.2</td>
<td>0.2</td>
<td>1</td>
<td>.001</td>
<td>.002</td>
<td>.004</td>
<td>.003</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>.000</td>
<td>.001</td>
<td>.094</td>
<td>.006</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>.082</td>
<td>.005</td>
<td>.000</td>
<td>.000</td>
<td></td>
</tr>
</tbody>
</table>
Table 7: Estimated Type I error probabilities, non-overlapping, $n = 30$, $\beta = 1$, $\rho(X_1, X_2) = 0$

<table>
<thead>
<tr>
<th>g</th>
<th>h</th>
<th>VP</th>
<th>ZF</th>
<th>ZM2</th>
<th>ZF</th>
<th>ZM2</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0</td>
<td>0.0</td>
<td>1</td>
<td>.007</td>
<td>.027</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>.098</td>
<td>.039</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>.000</td>
<td>.025</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.0</td>
<td>0.2</td>
<td>1</td>
<td>.045</td>
<td>.030</td>
<td>.038</td>
<td>.033</td>
</tr>
<tr>
<td>2</td>
<td>.118</td>
<td>.039</td>
<td>.213</td>
<td>.047</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>.021</td>
<td>.031</td>
<td>.013</td>
<td>.031</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.2</td>
<td>0.0</td>
<td>1</td>
<td>.009</td>
<td>.026</td>
<td>.010</td>
<td>.026</td>
</tr>
<tr>
<td>2</td>
<td>.091</td>
<td>.038</td>
<td>.091</td>
<td>.036</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>.000</td>
<td>.032</td>
<td>.000</td>
<td>.028</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.2</td>
<td>0.2</td>
<td>1</td>
<td>.052</td>
<td>.033</td>
<td>.047</td>
<td>.028</td>
</tr>
<tr>
<td>2</td>
<td>.123</td>
<td>.037</td>
<td>.213</td>
<td>.047</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>.021</td>
<td>.034</td>
<td>.016</td>
<td>.039</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

XN: X has bivariate normal distribution
XGH: X has a bivariate g-and-h distribution
ZF: Zou’s method with Fisher’s z
ZM2 is Zou’s method with M2.

Table 8: Estimated Type I error probabilities, non-overlapping, $n = 30$, $\beta = 1$, $\rho(X_1, X_2) = 0.5$

<table>
<thead>
<tr>
<th>g</th>
<th>h</th>
<th>VP</th>
<th>ZF</th>
<th>ZM2</th>
<th>ZF</th>
<th>ZM2</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0</td>
<td>0.0</td>
<td>1</td>
<td>.005</td>
<td>.030</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>.098</td>
<td>.038</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>.001</td>
<td>.031</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.0</td>
<td>0.2</td>
<td>1</td>
<td>.039</td>
<td>.024</td>
<td>.025</td>
<td>.036</td>
</tr>
<tr>
<td>2</td>
<td>.119</td>
<td>.032</td>
<td>.202</td>
<td>.041</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>.011</td>
<td>.042</td>
<td>.006</td>
<td>.045</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.2</td>
<td>0.0</td>
<td>1</td>
<td>.004</td>
<td>.020</td>
<td>.009</td>
<td>.030</td>
</tr>
<tr>
<td>2</td>
<td>.092</td>
<td>.032</td>
<td>.100</td>
<td>.031</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>.000</td>
<td>.039</td>
<td>.000</td>
<td>.033</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.2</td>
<td>0.2</td>
<td>1</td>
<td>.051</td>
<td>.026</td>
<td>.040</td>
<td>.037</td>
</tr>
<tr>
<td>2</td>
<td>.120</td>
<td>.034</td>
<td>.204</td>
<td>.040</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>.017</td>
<td>.053</td>
<td>.010</td>
<td>.051</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 9: Estimated Type I error probabilities, non-overlapping, $n = 30$, $\beta = 0$, $\rho(X_1, X_2) = 0.0$

<table>
<thead>
<tr>
<th>g</th>
<th>h</th>
<th>VP</th>
<th>ZF</th>
<th>ZM2</th>
<th>ZF</th>
<th>ZM2</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0</td>
<td>0.0</td>
<td>1</td>
<td>.045</td>
<td>.027</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>.075</td>
<td>.034</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>.004</td>
<td>.022</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.0</td>
<td>0.2</td>
<td>1</td>
<td>.039</td>
<td>.027</td>
<td>.049</td>
<td>.027</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>.123</td>
<td>.031</td>
<td>.252</td>
<td>.037</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>.011</td>
<td>.029</td>
<td>.002</td>
<td>.018</td>
</tr>
<tr>
<td>0.2</td>
<td>0.0</td>
<td>1</td>
<td>.03</td>
<td>.020</td>
<td>.035</td>
<td>.025</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>.083</td>
<td>.025</td>
<td>.079</td>
<td>.031</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>.000</td>
<td>.037</td>
<td>.000</td>
<td>.031</td>
</tr>
<tr>
<td>0.2</td>
<td>0.2</td>
<td>1</td>
<td>.045</td>
<td>.030</td>
<td>.041</td>
<td>.023</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>.118</td>
<td>.031</td>
<td>.185</td>
<td>.035</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>.019</td>
<td>.046</td>
<td>.012</td>
<td>.051</td>
</tr>
</tbody>
</table>

Table 10: Estimated Type I error probabilities, non-overlapping, $n = 30$, $\beta = 0$, $\rho(X_1, X_2) = 0.5$

<table>
<thead>
<tr>
<th>g</th>
<th>h</th>
<th>VP</th>
<th>ZF</th>
<th>ZM2</th>
<th>ZF</th>
<th>ZM2</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0</td>
<td>0.0</td>
<td>1</td>
<td>.045</td>
<td>.029</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>.090</td>
<td>.036</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>.000</td>
<td>.038</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.0</td>
<td>0.2</td>
<td>1</td>
<td>.045</td>
<td>.035</td>
<td>.045</td>
<td>.039</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>.132</td>
<td>.036</td>
<td>.245</td>
<td>.033</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>.010</td>
<td>.029</td>
<td>.002</td>
<td>.028</td>
</tr>
<tr>
<td>0.2</td>
<td>0.0</td>
<td>1</td>
<td>.042</td>
<td>.038</td>
<td>.048</td>
<td>.029</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>.108</td>
<td>.030</td>
<td>.161</td>
<td>.038</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>.002</td>
<td>.036</td>
<td>.005</td>
<td>.022</td>
</tr>
<tr>
<td>0.2</td>
<td>0.2</td>
<td>1</td>
<td>.042</td>
<td>.028</td>
<td>.051</td>
<td>.037</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>.132</td>
<td>.039</td>
<td>.257</td>
<td>.040</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>.015</td>
<td>.037</td>
<td>.002</td>
<td>.018</td>
</tr>
</tbody>
</table>