7-17-2020

Regression: Determining Which of p Independent Variables Has the Largest or Smallest Correlation With the Dependent Variable, Plus Results on Ordering the Correlations Winsorized

Rand Wilcox
University of Southern California, rwilcox@usc.edu

INVITED ARTICLE

Regression: Determining Which of *p* Independent Variables Has the Largest or Smallest Correlation with the Dependent Variable, Plus Results on Ordering the Correlations Winsorized

Rand Wilcox
University of Southern California
Los Angeles, CA

In a regression context, consider *p* independent variables and a single dependent variable. The paper addresses two goals. The first is to determine the extent it is reasonable to make a decision about whether the largest estimate of the Winsorized correlations corresponds to the independent variable that has the largest population Winsorized correlation. The second is to determine the extent it is reasonable to decide that the order of the estimates of the Winsorized correlations correctly reflects the true ordering. Both goals are addressed by testing relevant hypotheses. Results in Wilcox (in press) suggest using a multiple comparisons procedure designed specifically for the situations just described, but execution time can be quite high. A modified method for dealing with this issue is proposed.

Keywords: multiple comparisons, familywise error, robust methods, Winsorized correlation, ranking and selection

Introduction

Consider a situation involving some dependent variable, *Y*, and *p* covariates or independent variables $X_1, \ldots, X_p$. Let τ_j $(j = 1, \ldots, p)$ be some measure of the strength of the association between X_j and *Y* and let $\tau_{(1)} \leq \cdots \leq \tau_{(p)}$ denote these measures of association written in ascending order. Let $\hat{\tau}_{(j)}$ be some estimate of τ_j and denote the ordered estimates by $\hat{\tau}_{(1)} \leq \cdots \leq \hat{\tau}_{(p)}$. The primary focus in this paper

is on determining the extent it is reasonable to conclude that the independent variable corresponding to $\hat{\tau}_{(p)}$ is indeed the variable corresponding to $\tau_{(p)}$, the variable with the highest population measure of association. There are two components to this problem. The first is developing a reasonable decision rule. Second, given a decision rule, there is the issue of characterizing the probability of making a decision about whether $\hat{\tau}_{(p)}$ corresponds to $\tau_{(p)}$ as well as the probability of making a correct decision given that a decision is made. A related goal is to determine whether it is reasonable to decide that that $\hat{\tau}_{(1)}$ corresponds to $\tau_{(1)}$. Another goal is to whether it is reasonable to decide that $\hat{\tau}_{(1)} \leq \cdots \leq \hat{\tau}_{(p)}$ reflects the true ordering of the population measures of association. The focus here is on the Winsorized correlation, but the method to be described is readily extended to other robust measures of association.

First consider the goal of determining whether the correlation with the largest estimate does indeed correspond to the independent variable with the largest population correlation. A trivial modification of the approach used here can be used to deal with the situation where the goal is to make a decision about which independent variable has the smallest correlation instead. The decision rule used here is based on testing

$$H_0 : \tau_{\pi(j)} = \tau_{\pi(p)} \tag{1}$$

for each $j = 1, \ldots, p - 1$, where $\tau_{\pi(j)}$ is the measure of association associated with the group having the jth largest estimate. That is, compare the strength of the association of the covariate having the highest estimate to the strength of the association associated with each of the remaining $p - 1$ covariates. This is done in a manner that controls the familywise error (FWE) rate, meaning the probability of making one or more Type I errors, when $\tau_{(1)} \leq \cdots \leq \tau_{(p)}$. If all $p - 1$ hypotheses are rejected, decide the covariate yielding the largest estimate, $\hat{\tau}_{(p)}$, is in fact the covariate with the largest measure of association. In terms of making a decision about the order of the Winsorized correlations, the approach is to test

$$H_0 : \tau_{\pi(j)} = \tau_{\pi(j+1)} \tag{2}$$

($j = 1, \ldots, p - 1$). If each of these $p - 1$ hypotheses is rejected, decide that $\hat{\tau}_{(1)} \leq \cdots \leq \hat{\tau}_{(p)}$ reflects the true ordering.

When testing (1) or (2), the sign of the correlations plays a role. That is, a correlation equal to -0.6 is considered to be less than a correlation of 0.5 even though the strength of the association reflected by the first correlation is estimated

3

to be stronger. If there is compelling evidence that an independent variable has a negative association with the dependent variable, one could simply multiply this independent variable by -1 to make the correlation positive and then test the relevant hypotheses. If, for example, a correlation of -0.6 is converted to a correlation of 0.6, and if now this variable is significantly larger than the other correlations, decide that the variable initially having a correlation of -0.6 has the strongest association with Y.

Regarding the goal of controlling the familywise error (FWE) rate, a seemingly simple solution is to perform each test using the bootstrap method in Wilcox (2017a, section 11.10.1) and then use some improvement on the Bonferroni method to control the FWE rate (e.g., Hochberg, 1988; Hommel, 1988). However, this approach is unsatisfactory for reasons illustrated in the next section. Also, the main goal is not to control the FWE rate, but rather characterize the probability of making a correct decision when a decision is made.

Comparing Winsorized Correlations

First, the method for computing a Winsorized correlation is reviewed. Consider a single independent variable and let (Y_i, X_i) $(i = 1, \ldots, n)$ denote a random sample. Let $Y_{(1)} \leq \cdots \leq Y_{(n)}$ denote the Y values written in ascending order. Let $g = [0.2n]$, where $[0.2n]$ is $0.2n$ rounded down to the nearest integer. The Winsorized values of the dependent variable are $W_i = Y_{(g+1)}$ if $Y_i \leq Y_{(g+1)}$, $W_i = Y_i$ if $Y_{(g+1)} < Y_i < Y_{(n-g)}$, and $W_i = Y_{(n-g)}$ if $Y_i \geq Y_{(n-g)}$. The Winsorized values of the independent variable X are computed in a similar fashion yielding say $U_1, \ldots, U_n$. The Winsorized correlation between X and Y is just Pearson's correlation based on (W_i, U_i), $i = 1, \ldots, n$.

Now consider $p = 2$ where the goal is test $H_0 : \tau_1 = \tau_2$, the hypothesis that two independent variables have the same Winsorized correlation with Y. A basic percentile bootstrap method has been found to perform well (Wilcox, 2017), which is applied as follows:

1.   Generate a bootstrap sample by randomly sampling with replacement n vectors of values from $(Y_1, X_{11}, X_{12}), \ldots, (Y_n, X_{n1}, X_{n2})$ yielding $(Y_1^*, X_{11}^*, X_{12}^*), \ldots, (Y_n^*, X_{n1}^*, X_{n2}^*)$.

2.   Compute the Winsorized correlation between Y and the jth independent variable based on this bootstrap sample yielding $\hat{\tau}_1^*$ and $\hat{\tau}_2^*$ and let $d^* = \hat{\tau}_1^* - \hat{\tau}_2^*$.

3.   Repeat steps 1 and 2 B times and let d_b^* ($b = 1, \ldots, B$) denote the resulting d^* values.

4.   Put the $d_1^*, \ldots, d_B^*$ values in ascending order and label the results $d_{(1)}^* \leq \cdots \leq d_{(B)}^*$.

5.   Let $\ell = \alpha B/2$, rounded to the nearest integer and $u = B - 1$. Then the $1 - \alpha$ confidence interval for $\tau_1 - \tau_2$ is $(d_{(\ell+1)}^*, d_{(u)}^*)$.

6.   Let $A = \sum I(d_b^*)$, where the indicator function $I(d_b^*) = 1$ if $d_b^* < 0$; otherwise $I(d_b^*) = 0$.

Letting $P^* = A/B$, a p-value is

$$P = 2\min\{P^*, 1 - P^*\} \text{ (Liu \& Singh, 1997)} \tag{3}$$

When testing (1) or (2), p-values are computed as just described, only now bootstrap samples are generated by resampling with replacement from $(Y_1, X_{11}, \ldots, X_{1p}), \ldots, (Y_n, X_{n1}, \ldots, X_{np})$. Once the p-values have been computed, a seemingly simple approach to controlling the FWE rate is to use the well-known Bonferroni method. That is, test each of the $p - 1$ hypotheses at the $\alpha/(p - 1)$ level with the goal of ensuring that the actual FWE rate is less than 0.05.

However, this approach needs to be adjusted. To illustrate why, suppose the jth hypothesis is rejected when $P_j \leq c_j$, where P_j is the p-value associated with the jth hypothesis. It is informative and useful to determine $c_1, \ldots, c_{p-1}$ for a special case: $p = 4$, $n = 50$, and where all $p + 1$ random variables have independent standard normal distributions. A simulation was performed to estimate the distribution of the $p - 1$ p-values based on $L = 3000$ replications and $B = 500$ bootstrap samples. Let $\hat{P}$ denote the resulting matrix of p-values having L rows and $p - 1$ columns. Then the jth column of $\hat{P}$ provides an estimate of c_j. If the goal is to have a Type I error probability equal to α, the estimate of c_j is obtained via some quantile estimator applied to the jth column of $\hat{P}$. Here, the Harrell and Davis (1982) estimator is used. For the situation at hand, the estimates of the $\alpha = 0.05$ quantiles based on the columns of $\hat{P}$ are $(\hat{c}_1, \hat{c}_2, \hat{c}_3) = (0.1961, 0.0727, 0.0136)$.

Now consider the strategy of using the Bonferroni method to control the FWE rate. The Bonferroni method assumes that for each individual test, the probability of a Type I error is α. That is, if a hypothesis is rejected at the α level, this corresponds to rejecting if the p-value is less than or equal to α. As just illustrated, this is not remotely accurate for the situation at hand. In terms of achieving a FWE rate less than or equal to 0.05, the result $(\hat{c}_1, \hat{c}_2, \hat{c}_3) = (0.1961, 0.0727, 0.0136)$

reflects a practical concern, especially when dealing with a large sample size: The actual FWE rate can be substantially higher than the nominal level.

As an illustration, consider the case $p = 4$ and where all $p + 1$ variables have a standard normal distribution with a common correlation of 0.5. Using $B = 500$ bootstrap samples, a simulation based on 1000 replications estimated the actual FWE rate to be 0.066, 0.075 and 0.086 for sample sizes 40, 100 and 300, respectively. For $p = 5$, the estimates were 0.077 and 0.11 for sample sizes 40 and 100, respectively. By implication, improvements on the Bonferroni method previously mentioned are unsatisfactory as well because they reject as many or more hypotheses at the Bonferroni method.

An outline of a method for controlling the FWE rate is as follows. Momentarily assume that the $p + 1$ variables have a multivariate normal distribution with a common correlation equal to zero and use a simulation to estimate the joint distribution of the $p - 1$ p-values yielding $\hat{P}$ as just described. Note that $\hat{P}$ can be used to compute a corrected p-value for jth hypothesis:

$$\tilde{P}_j = \frac{1}{L} \sum I\left(P_j \geq \hat{P}_{ij}\right),$$

where P_j is the p-value when testing the jth hypothesis, $\hat{P}_{ij}$ is the element in the ith row and jth column of $\hat{P}$ and the indicator function $I\left(P_j \geq \hat{P}_{ij}\right) = 1$ if $P_j \geq \hat{P}_{ij}$, otherwise $I\left(P_j \geq \hat{P}_{ij}\right) = 0$. Next, use Hochberg's improvement on the Bonferroni method based on the adjusted p-values, which is applied as follows. Put the $p - 1$ p-values in descending order and label the results $\tilde{P}_{[1]} \geq \cdots \geq \tilde{P}_{[p-1]}$. Set $k = 1$ and reject all $p - 1$ hypotheses if

$$\tilde{P}_{[k]} \leq \alpha / k.$$

If $\tilde{P}_{[1]} >$, proceed as follows:

1.   Increment k by 1. If

$$\tilde{P}_{[k]} \leq \alpha / k,$$

6

2. stop and reject all hypotheses having a p-value less than or equal to $\tilde{P}_{[k]}$.

3. If $\tilde{P}_{[k]} > \alpha / k$, repeat steps 1 and 2 until a significant result is obtained or all $p - 1$ hypotheses have been tested. This will be called method S henceforth.

Two practical issues remain. The first is that computing $\hat{P}$ can require several minutes of execution time when using the R function mentioned in the final section of this paper. For example, with $p = 4$, $n = 100$, $B = 500$ and $L = 1000$, execution time on a MacBook Pro with a 2.9GHz processor was a little over 14 minutes. To deal with this, consideration is given to using an estimate of P based $n = 50$. The idea is that the estimated quantiles of the p-value distributions, based on the resulting estimate of $\hat{P}$, could be stored in appropriate software and used with any sample size. This will be called method FS. To the extent this approach controls the FWE rate reasonably well, low execution time is achieved.

Now an issue is whether these quantiles continue to perform reasonably well as n increases. The second major issue is the impact of non-normality as well as situations where there is an association among all of the $p + 1$ variables. Simulations results reported in the next section deal these issues.

Note that both methods S and FS are readily modified to testing (2), which will be called methods O and FO, respectively.

Simulation

Simulations were used to assess the actual FWE rate using method FS. Data were generated from one of four distributions: normal, symmetric and heavy-tailed, skewed and light-tailed, and skewed and heavy-tailed. More precisely, data were generated from g-and-h distributions (Hoaglin, 1985), which arise as follows. Let Z be a random variable having a standard normal distribution. Then

$$\frac{\exp(gZ) - 1}{g} \exp(hZ^2 / 2),$$

if $g > 0$ and

$$Z \exp\left(hZ^2 / 2 \right)$$

if $g = 0$ has a g-and-h distribution, where g and h are parameters that determine the first four moments. The four distributions used here are the standard normal ($g = h = 0$), a symmetric heavy-tailed distribution ($h = 0.2, g = 0$), an asymmetric distribution with relatively light tails ($h = 0, g = 0.2$), and an asymmetric distribution with heavy tails ($g = h = 0.2$). Table 1 summarizes the skewness (κ_1) and kurtosis (κ_2) of these distributions.

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

g	h	κ_1	κ_2
0.0	0.0	0.00	3.00
0.0	0.2	0.00	21.46
0.2	0.0	0.61	3.68
0.2	0.2	2.81	155.98

Two choices for a common correlation among all $p + 1$ variables were used: $\rho = 0.0$ and $\rho = 0.5$. For $\rho = 0.5$, first data were generated data from a multivariate normal distribution and then the marginal distributions were transformed to g-and-h distributions. Transforming to a g-and-h distribution can alter somewhat the correlation among the covariates. However, this is easily corrected using the R function rngh in Wilcox (2017, section 4.2.1).

The simulation results for method FS are shown in Table 2 and based on 1000 replications. Although the seriousness of a Type I error can depend on the situation, Bradley (1978) suggests that as a general guide, when testing at the 0.05 level, the actual level should be between 0.025 and 0.075. As can be seen, this criterion is met for all of the situations considered. For $p = 4$, the estimates range between 0.037 and 0.044. For $p = 6$, there is more variation, the estimates ranging between 0.026 and 0.067. All indications are that the approximation of the null distribution of the p-values when $n = 50$ provides a good approximation of the null distributions for sample sizes between 40 and 400.

Some simulations were run where the goal was to identify the independent variable having the lowest correlation. This was done by testing

$$H_0 : \tau_{\pi(j)} = \tau_{\pi(1)}, \tag{4}$$

for each j, $(j = 2, \ldots, p)$. The same critical values used by method FS were used here. The results were virtually the same as those in Table 2.

Table 2. Estimates of α, $\hat{\alpha}$, when testing at the 0.05 level using method FS

			$p = 4$		$p = 6$	
g	h	ρ	$n = 40$	$n = 400$	$n = 40$	$n = 400$
0.0	0.0	0.0	0.042	0.044	0.030	0.067
0.0	0.2	0.0	0.038	0.042	0.029	0.065
0.2	0.0	0.0	0.039	0.044	0.032	0.066
0.2	0.2	0.0	0.037	0.043	0.026	0.062
0.0	0.0	0.5	0.042	0.044	0.033	0.067
0.0	0.2	0.5	0.038	0.042	0.029	0.065
0.2	0.0	0.5	0.039	0.044	0.032	0.066
0.2	0.2	0.5	0.037	0.043	0.026	0.062

Results for method FO, where the goal is to test (2), are shown in Table 3. As can be seen, control over the Type I error probability meets Bradley's criterion when $p = 4$. For $p = 6$, method FO performs well for $n = 40$, but for $n = 400$, estimates exceed 0.075, the largest being 0.080. Lowering n to 350, not shown in Table 3, method FO performs reasonably well. For $g = h = 0$ the estimate was 0.069. For $n = 300$, now the estimate was 0.055.

Table 3. Estimates of α, $\hat{\alpha}$, when testing at the 0.05 level using method FO

			$p = 4$		$p = 6$	
g	h	ρ	$n = 40$	$n = 400$	$n = 40$	$n = 400$
0.0	0.0	0.0	0.061	0.065	0.050	0.073
0.0	0.2	0.0	0.059	0.062	0.047	0.070
0.2	0.0	0.0	0.062	0.066	0.042	0.080
0.2	0.2	0.0	0.058	0.063	0.044	0.077
0.0	0.0	0.5	0.061	0.065	0.050	0.073
0.0	0.2	0.5	0.059	0.062	0.047	0.070
0.2	0.0	0.5	0.062	0.066	0.042	0.080
0.2	0.2	0.5	0.058	0.063	0.044	0.077

A Ranking and Selection Perspective and a Limitation of Methods F and FS

Methods F and FS are designed to control the FWE rate when $\tau_{(1)} = \cdots = \tau_{(p)}$. However, consider the situation where $p = 5$, $n = 80$ and the goal is to have a FWE rate equal to 0.05. The critical p-values based on method F are $(\hat{c}_1, \hat{c}_2, \hat{c}_3, \hat{c}_4) = (0.232, 0.088, 0.036, 0.004)$. Now suppose $\tau_{(p-1)} = \tau_{(p)}$ and that $\tau_{(p)}$ is substantially larger than $\tau_{(1)} \ldots \tau_{(p-2)}$. That is, with near certainty, $H_0 : \tau_{\pi(p-1)} = \tau_{\pi(p)}$, the only true hypothesis, will be tested at the 0.232 level resulting in a Type I error greater than α for the one situation where two independent variables have the same Winsorized correlation with the dependent variable. Simulations confirm that for this particular situation, testing at the 0.05 level is more appropriate.

However, there is an alternative perspective that might be deemed useful: view the problem in a manner similar to the literature dealing with ranking and selection techniques (e.g., Bechhofer et al., 1968; Gibbons et al., 1987; Gupta & Panchapakesan, 1987; Mukhopadhyay & Solanky, 1994). To describe what this means, first note that from basic principles, when there is independence, the Winsorized correlation, as well other robust measures of association, are equal to zero. However, when there is an association, the basic argument made by Tukey (1991) extends to the situation at hand: surely any two Winsorized correlations differ at some decimal place.

What is needed is some rule for making a decision about which variable has the strongest association and then finding some useful measure of how well the method is performing. Mimicking the approach used in the ranking and selection literature in an obvious way, let $\delta = \tau_{(p)} - \tau_{(p-1)}$. If δ is small, deciding that the independent variable associated with $\tau_{(p-1)}$ has the strongest association is not that important. But if $\delta \geq \delta^*$, for some specified value for δ^*, deciding that the independent variable associated $\tau_{(p)}$ has the larger Winsorized correlation is important. For example, Cohen (1988) suggested that as a general guide, Pearson correlations 0.1, 0.3 and 0.5 be considered as small, medium and large, respectively. When $\tau_{(1)} = \cdots = \tau_{(p-1)} = 0$, for instance, attention might be focused on the extent a correct decision is made for δ^* equal to 0.1, 0.3 and 0.5. In the context of method F, what is the probability of a correct decision given that a decision has been made regarding which independent variable has the highest correlation? Noting that the performance of method FS does not appear to be overly sensitive to which distribution generated the data, this probability is readily estimated via a simulation, given n and p. Briefly, generate data where the marginal distributions are standard

normal, $\tau_{(1)} = \cdots = \tau_{(p-1)} = 0$ and $\tau_{(p)} = \delta^*$. If a decision is made using method FS about which independent variable has the highest correlation, note whether a correct decision was made. Repeating this process many times, the proportion of correct decisions among the situations where a decision is made estimates the true probability of a correct decision. In all likelihood the actual probability is even higher when $\delta = \delta^*$ because based on Tukey's argument, surely $\tau_{(1)} < \cdots < \tau_{(p-1)}$. In a slightly broader context, one might focus on $\tau_{(1)} = \cdots = \tau_{(p-1)} = \tau^* > 0$, say, and $\tau_{(p)} = \tau^* + \delta^*$.

Illustration

The proposed methods are illustrated using data from the Well Elderly II study (Clark et al., 2011), which pertained to improving the physical and mental well-being of older adults. The sample size is $n = 232$. The dependent variable is taken to be a measure of depressive symptoms (CESD) before intervention. Here, the focus is on the strength of the association of CESD with four other measures: the change in cortisol measures taken upon awakening and measure again 30-45 minutes later (CAR), meaningful activities (MAPA), stress and a measure of life satisfaction (LSIZ). The estimates of the Winsorized correlations for CAR, MAPA, stress and LSIZ were 0.0162, −0.46, 0.61 and −0.53, respectively. The latter three of these variables had a significant association with CESD with p-values less than 0.001. Multiplying both MAPA and LSIZ by −1 and applying method FS, no decision was made about which independent variable had the largest association. However, a decision was made about CAR having the lowest correlation. All three of the Hochberg adjusted p-values were less than or equal to 0.001.

Given a decision was not made about which variable had the largest Winsorized correlation, it is of interest to characterize the likelihood of making a decision. Based on $n = 232$, $p = 4$, $\tau_{(1)} = \tau_{(2)} = \tau_{(3)} = 0$ and $\tau_{(4)} = 0.1$, the probability of making a decision was estimated to 0.06 based on a simulation with 500 replications. The 0.95 confidence interval for this probability is (0.042, 0.085). For $\tau_{(4)} = 0.3$, the estimate is now 0.764 and the 0.95 confidence interval is (0.725, 0.799). Because a decision was made about the lowest Winsorized correlation, there is issue of characterizing the likelihood that the decision is correct. The probability of making a correct decision when $\tau_{(1)} = \tau_{(2)} = \tau_{(3)} = 0.1$ and $\tau_{(4)} = 0$ was estimated to be and 0.918; the 0.95 confidence interval is (0.803, 0.973).

Conclusion

All indications are that the strategy for reducing execution time, used by method FS, performs well when the goal is to determine whether it is reasonable to make a decision about which independent variable has the strongest or smallest Winsorized correlation with the dependent variable. This remains the case when the focus is on ordering the Winsorized correlations (method FO) by testing (2) provided that the sample size is no larger than 350. Otherwise, it is safer to use method O rather than method FO. That is, use an estimate of P based on the sample size stemming from the study, which can result in high execution time.

Method FS is designed to control the FWE rate when all $p - 1$ of the hypotheses given by (1) are true. But situations can be constructed where a Type I error rate can exceed the nominal level. The same is true in related methods where the goal is to identify which of J independent groups has the largest measure of location or the highest probability of success (Wilcox, in press; Wilcox, 2019). As was the case here, it is a simple matter to estimate the probability of a correct decision, given that a decision is made, within the context of an indifference zone. It remains unknown how these methods might be modified so that the Type I error rate never exceeds the nominal level.

A few simulations were run with the Winsorized correlation replaced by Spearman's rho and Kendall's tau. For method FS, control over the FWE rate was found to satisfy Bradley's criterion among the situations considered. For $p = 4$, $g = h = 0$, and $n = 40$, the estimate of the FWE rate using Spearman's rho was 0.057 and it was 0.062 using Kendall's tau. For the $p = 6$ the estimates were 0.065 and 0.055, respectively. But a more comprehensive simulation is needed. A more cautious approach, for the moment, is to use methods F and O when using Spearman's rho or Kendall's tau.

An open issue is how methods S and O might be extended to Pearson's correlation. There are methods for testing (1) (e.g., Wilcox, 2017, section 11.10.1). But even when there are only two independent variables, if all three variables have a reasonably strong association, the actual Type I error probability can be well below the nominal level, even under normality and homoscedasticity. For the situation at hand, some improved method for comparing Pearson's correlations is needed.

The R function `corCOM.DVvsIV` applies method FS by default and is stored in the file Rallfun-v37 located at https://dornsife.usc.edu/cf/labs/wilcox/wilcox-faculty-display.cfm as well as https://osf.io/nvd59/quickfiles. To use method S, set the argument `com.p.dist = TRUE`. When using method S, if the same sample size

will be used in future analyses, the estimate of $\boldsymbol{P}$ can be computed via the R function `corCOM.DVvsIV.crit` and stored in some R variable, say A. Then, when using `corCOM.DVvsIV`, set the argument `PV = A`. The argument `corfun` controls which correlation is used. The default is a 20% Winsorized correlation. Setting `corfun = tau` would use Kendall's tau and `corfun = spear` would use Spearman's rho. When `com.p.dist = TRUE`, the argument `iter` indicates the number of replications, L, used to compute $\hat{\boldsymbol{P}}$ and the argument `nboot` indicates B, the number of bootstrap samples. The R function `corREGorder` is exactly like the function `corCOM.DVvsIV` only now the goal is to use method O or FO. The default is to use method FO. Similar to method S, $\hat{\boldsymbol{P}}$ can be computed via the R function `corREGorder.crit`. The result can be passed to `corREGorder` via the argument `PV`. Based on the indifference zone perspective, the R function `corCOM.PMDPCD` estimates the probability of a correct decision given that a decision has been made using method FS. By default, when the focus is on which independent variable has the largest measure of association, this is done for $\tau_{(1)} = \cdots = \tau_{(p-1)} = 0$ and $\tau_{(p)} = \delta^*$, where δ^* is specified by the argument `delta`. By default, `delta = 0.3` is used. When the goal is to identify the variable with the smallest correlation, now the function uses $\tau_{(1)} = \cdots = \tau_{(p-1)} = \delta^*$, and $\tau_{(p)} = 0$.

References

Bechhofer, R.E., Kiefer, J. & Sobel, M. (1968). *Sequential Identification and Ranking Procedures*. Chicago, Ill: University of Chicago Press

Bradley, J. V. (1978) Robustness? *British Journal of Mathematical and Statistical Psychology, 31*(2), 144–152. https://doi.org/10.1111/j.2044-8317.1978.tb00581.x

Clark, F., Jackson, J., Carlson, M., Chou, C.-P., Cherry, B. J., Jordan-Marsh, M., Knight, B. G., Mandel, D. Blanchard, J., Granger, D. A., Wilcox, R. R., Lai, M. Y., White, B., Hay, J., Lam, C., Marterella, A., & Azen, S. P. (2011). Effectiveness of a lifestyle intervention in promoting the well-being of independently living older people: results of the Well Elderly 2 Randomise Controlled Trial. *Journal of Epidemiology and Community Health, 66*(9), 782–790. https://doi.org/10.1136/jech.2009.099754

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences.* (2nd Ed.). Hillsdale, NJ: Erlbaum. https://doi.org/10.4324/9780203771587

Gibbons, J., Olkin, I. and Sobel, M. (1987). *Selecting and Ordering Populations: A New Statistical Methodology*. (2nd. Ed.). New York: Wiley.

Gupta, S. S. & Panchapakesan, S. (1987). Multiple Decision Procedures: Theory and Methodology of Selecting and Ranking Populations. Philadelphia, PA: SIAM.

Harrell, F. E. & Davis, C. E. (1982). A new distribution-free quantile estimator. *Biometrika, 69*(3), 635–640. https://doi.org/10.1093/biomet/69.3.635

Hoaglin, D. C. (1985). Summarizing shape numerically: The g-and-h distribution. In D. Hoaglin, F. Mosteller & J. Tukey (Eds.) Exploring Data Tables Trends and Shapes. New York: Wiley, pp. 461–515.

Hommel, G. (1988). A stagewise rejective multiple test procedure based on a modified Bonferroni test. *Biometrika, 75*(2), 383–386. https://doi.org/10.1093/biomet/75.2.383

Hochberg, Y. (1988). A sharper Bonferroni procedure for multiple tests of significance. *Biometrika, 75*(4), 800–802. https://doi.org/10.1093/biomet/75.4.800

Liu, R. G. & Singh, K. (1997). Notions of limiting P values based on data depth and bootstrap. *Journal of the American Statistical Association, 92*(437), 266–277. https://doi.org/10.2307/2291471

Mukhopadhyay, N. & Solanky. T. (1994). Multistage Selection and Ranking Procedures: Second Order Asymptotics. New York: Marcel Dekker.

Tukey, J. W. (1991). The philosophy of multiple comparisons. *Statistical Science, 6*(1), 100-116. https://doi.org/10.1214/ss/1177011945

Wilcox, R. R. (2017). *Introduction to Robust Estimation and Hypothesis Testing*. 4th Edition. San Diego, CA: Academic Press

Wilcox, R. R. (in press). Inferences about which of *J* independent groups has the largest robust measure of location. *Journal of Modern Applied Statistical Methods*.

Wilcox, R. R. (2019). Inferences about which of *J* independent binomial distributions has the largest probability of success. *Journal of Modern Applied Statistical Methods*, 18(2), eP3359. https://doi.org/10.22237/jmasm/1604190960