Supplementary Material for Stable Combination Tests

Xing Ling and Yeonwoo Rho

Michigan Technological University

This supplementary material presents a proof for Theorem 2 in Section S1 and our simulation setting and results in Section S2.

S1 Proof for Theorem 2

This section presents a proof for Theorem 2. We first present three lemmas. Recall that $p(x) = 2 - 2\Phi(|x|)$ for all $x \in \mathbb{R}$. Lemmas 1 and 2 help find the lower bounds of $F^{-1}(1 - \min_{i\in S} p_i)$ and $F^{-1}(1 - \max_{i\in S} p_i)$, respectively. Lemma 3 presents a lower bound for $a_{n;\alpha}$.

Lemma 1. *Let* $g(x) = c_{\alpha,\beta} x^{1/\alpha} e^{x^2/(2\alpha)}$ *where* $c_{\alpha,\beta} = \left[\frac{1+\beta}{\sqrt{2\pi}}\Gamma(\alpha)\sin\left(\frac{\pi\alpha}{2}\right)\right]^{1/\alpha}$ *is a constant,* $0 < \alpha < 2$ *and* $-1 < \beta \leq 1$*. For* $x \to \infty$*,*

$$F^{-1}[1 - p(x)|\alpha, \beta] > g(x).$$

Proof of Lemma 1. When $x \to \infty$, $g(x) \to \infty$. Therefore, we can apply the

Corresponding author: Yeonwoo Rho, Department of Mathematical Sciences, Michigan Technological University, Houghton, MI, 49931, USA. E-mail: yrho@mtu.edu.

right tail approximation of a stable distribution in Theorem 1.2 form Nolan (2020). When $0 < \alpha < 2$ and $-1 < \beta \leq 1$, for large enough x,

$$\begin{aligned} 1 - F[g(x)|\alpha, \beta] &= \Pr[W_{0;\alpha,\beta} > g(x)] \\ &\sim \frac{1+\beta}{\pi}\Gamma(\alpha)\sin\left(\frac{\pi\alpha}{2}\right)[g(x)]^{-\alpha} \\ &= \sqrt{\frac{2}{\pi}}x^{-1}e^{-x^2/2}. \end{aligned}$$

From Mill's ratio inequality that $1 - \Phi(x) \leq \frac{\phi(x)}{x}$ for any $x > 0$, where $\Phi(\cdot)$ and $\phi(\cdot)$ represent the distribution function and probability density function of a standard normal random variable respectively, we have

$$p(x) = 2[1 - \Phi(x)] \leq 2\frac{\phi(x)}{x} = \sqrt{\frac{2}{\pi}}x^{-1}e^{-x^2/2}.$$

Therefore, $p(x) \leq 1 - F[g(x)|\alpha, \beta]$ for large enough x. Since F^{-1} is increasing, $F^{-1}[1 - p(x)] > g(x)$ for large enough x. □

Lemma 2. *Define* $\tilde{g}(x) = -\tilde{c}_{\alpha,\beta}x^{-1/\min(\alpha,1)}e^{x^2/(2\alpha)}$ *with constant* $\tilde{c}_{\alpha,\beta} = \left[\frac{1-\beta}{\sqrt{2\pi}}\Gamma(\alpha)\sin\left(\frac{\pi\alpha}{2}\right)\right]^{1/\alpha}$ *where* $0 < \alpha < 2$ *and* $-1 \leq \beta \leq 1$*. When* $x \to 0^+$,

$$F^{-1}[1 - p(x)|\alpha, \beta] > \tilde{g}(x).$$

Proof of Lemma 2. We first consider the case where $-1 \leq \beta < 1$. Similarly to the proof of Lemma 1, when $x \to 0^+$, $\tilde{g}(x) \to -\infty$. Applying the left tail approximation from Theorem 1.2 of Nolan (2020) for $0 < \alpha < 2$ and

$-1 \leq \beta < 1$,

$$\begin{aligned} F[\tilde{g}(x)|\alpha,\beta] &= \Pr[W_{0;\alpha,\beta} < \tilde{g}(x)] \\ &\sim \frac{1-\beta}{\pi}\Gamma(\alpha)\sin\left(\frac{\pi\alpha}{2}\right)[-\tilde{g}(x)]^{-\alpha} \\ &= \sqrt{\frac{2}{\pi}}x^{\max(\alpha,1)}e^{-x^2/2}. \end{aligned}$$

In addition, the standard normal distribution function, $\Phi(x)$, can be rewritten using the integration by parts,

$$1 - p(x) = 2\Phi(x) - 1 = \sqrt{\frac{2}{\pi}}xe^{-x^2/2} + Q(x),$$

where $Q(x) = \sqrt{\frac{2}{\pi}}e^{-x^2/2}(\frac{x^3}{3} + \frac{x^5}{3*5} + \cdots) > 0$ for any $x > 0$. Noting that $x^{\max(\alpha,1)} \leq x$ for $x < 1$ and $\alpha > 1$, we conclude that $1 - p(x) > F[\tilde{g}(x)|\alpha,\beta]$ for $x \to 0^+$.

When $\beta = 1$, the distribution is totally skewed to the right, and the left tail probability does not follow a power law. Instead, we know that the left tail probability of $W_{0;\alpha,1}$ is smaller than that of $W_{0;\alpha,\beta}$ with $-1 \leq \beta < 1$. That is,

$$F[\tilde{g}(x)|\alpha,1] = \Pr[W_{0;\alpha,1} < \tilde{g}(x)] < \Pr[W_{0;\alpha,\beta} < \tilde{g}(x)].$$

Therefore, $1 - p(x) > F[\tilde{g}(x)|\alpha,\beta]$ for all $-1 \leq \beta \leq 1$. Since F^{-1} is increasing, we have $F^{-1}[1 - p(x)] > \tilde{g}(x)$, which completes the proof. □

Lemma 3. *Let $w_i \in (0,1)$ be nonnegative weights such that $\sum_{i=1}^n w_i = 1$. The normalizing constant $a_{n;\alpha} = (\sum_{i=1}^n w_i^\alpha)^{-1/\alpha} \geq \min(n^{1-1/\alpha}, 1)$.*

Proof of Lemma 3. The lower bound of $a_{n;\alpha}$ is considered in three separate cases. First, when $\alpha = 1$, $a_{n;\alpha} = 1$. The second case is when $0 < \alpha < 1$. By Hölder's inequality,

$$\sum_{i=1}^{n} w_i^{\alpha} \leq \left[\sum_{i=1}^{n} (w_i^{\alpha})^{1/\alpha}\right]^{\alpha} n^{1-\alpha} = n^{1-\alpha},$$

which is equivalent to $a_{n;\alpha} \geq n^{1-1/\alpha}$. The last case is when $1 < \alpha < 2$. Since the l_α norm is a decreasing function in α for any $\alpha \geq 1$,

$$\left(\sum_{i=1}^{n} |w_i|^{\alpha}\right)^{1/\alpha} \leq \sum_{i=1}^{n} |w_i| = 1,$$

and therefore, $a_{n;\alpha} \geq 1$. Combining the above three cases, we have $a_{n;\alpha} = (\sum_{i=1}^{n} w_i^{\alpha})^{-1/\alpha} \geq \min(n^{1-1/\alpha}, 1)$. □

Now we are ready to prove Theorem 2.

Proof of Theorem 2. Recall that the test statistic is defined as $T_{n;\alpha,\beta}(\boldsymbol{p}) = T_{n;\alpha,\beta}(\boldsymbol{X}) = a_{n;\alpha} \sum_{i=1}^{n} w_i F^{-1}[1-p(X_i)|\alpha,\beta]$, where $a_{n;\alpha} = \left(\sum_{j=1}^{n} w_j^{\alpha}\right)^{-1/\alpha}$. Under Assumption 4, the test statistic $T_{n;\alpha,\beta}(\boldsymbol{X})$ can be decomposed into two parts:

$$\begin{aligned} T_{n;\alpha,\beta}(\boldsymbol{X}) =& a_{n;\alpha} \sum_{i \in S} w_i F^{-1}[1 - p(X_i)|\alpha, \beta] \\ &+ a_{n;\alpha} \sum_{i \in S^c} w_i F^{-1}[1 - p(X_i)|\alpha, \beta] \\ :=& A_n + B_n. \end{aligned}$$

In order to show $T_{n;\alpha,\beta}(\boldsymbol{X}) \to \infty$ as $n \to \infty$, we will show that $A_n \to \infty$ with probability 1 and that B_n cannot be arbitrary large negative.

Part A_n can be further decomposed as follows:

$$\begin{aligned} A_n \geq & a_{n;\alpha} c_0 n^{-1} \max_{i \in S} F^{-1}[1 - p(X_i)|\alpha, \beta] \\ & + a_{n;\alpha}(\sum_{j \in S} w_j - c_0 n^{-1}) \min_{i \in S} F^{-1}[1 - p(X_i)|\alpha, \beta] \\ := & A_{n,1} + A_{n,2} \end{aligned}$$

In the following arguments, we will prove that $A_n \to \infty$ with probability 1 by showing that $A_{n,1}$ can be arbitrarily large whereas $A_{n,2} > o_p(1)$ as $n \to \infty$.

Since $F^{-1}[1 - p(x)|\alpha, \beta]$ is increasing in x, $A_{n,1} = a_{n;\alpha} c_0 n^{-1} F^{-1}[1 - p(\max_{i \in S} |X_i|)|\alpha, \beta]$. Recall that the set of positive signals, S_+, is assumed to have the cardinality no less than $|S|/2$. From Lemma 6 of Cai et al. (2014) and using the same argument as in the proof of Theorem 3 of Liu and Xie (2020), $\max_{i \in S} |X_i| \geq \mu_0 + \sqrt{2 \log |S_+|} + o_p(1)$. Given the assumptions that $\mu_0 = \sqrt{2r \log n}$ and $\sqrt{2 \log |S_+|} \geq \sqrt{2(\gamma \log n - \log 2)}$, we have $\max_{i \in S} |X_i| \to \infty$ with probability 1. Lemma 1 implies that

$$\Pr\left\{F^{-1}\left[1 - p\left(\max_{i \in S} |X_i|\right) \middle| \alpha, \beta\right] > g\left(\max_{i \in S} |X_i|\right)\right\} \underset{n \to \infty}{\longrightarrow} 1,$$

which is equivalent to

$$\Pr\left\{A_{n,1} \geq a_{n;\alpha} c_0 n^{-1} c_{\alpha,\beta} \left(\max_{i \in S} |X_i|\right)^{1/\alpha} \exp\left[\frac{(\max_{i \in S} |X_i|)^2}{2\alpha}\right]\right\} \underset{n \to \infty}{\longrightarrow} 1.$$

Noting that $\max_{i\in S}|X_i| \geq \sqrt{2r\log n}+\sqrt{2\log|S_+|}+o_p(1)$, $\Pr\{\max_{i\in S} X_i > 1\} \to 1$, and $\sqrt{\log|S_+|} \geq \sqrt{2(\gamma\log n - \log 2)} \approx \sqrt{2\gamma\log n}$, we have

$$\Pr\left\{A_{n,1} \geq a_{n;\alpha}c_0 n^{-1}c_{\alpha,\beta}\left[n^{(\sqrt{\gamma}+\sqrt{r})^2}\right]^{1/\alpha}\right\} \underset{n\to\infty}{\longrightarrow} 1.$$

From Lemma 3, $a_{n;\alpha} = \left(\sum_{j=1}^n w_j^\alpha\right)^{-1/\alpha} \geq \min(n^{(\alpha-1)/\alpha}, 1)$, and therefore

$$\Pr\left\{A_{n,1} \geq c_0 c_{\alpha,\beta}\left[n^{(\sqrt{\gamma}+\sqrt{r})^2/\alpha-1+\min(1-1/\alpha,0)}\right]\right\} \underset{n\to\infty}{\longrightarrow} 1.$$

According to Part 3 of Assumption 4, $\sqrt{r}+\sqrt{\gamma} > \max(\sqrt{\alpha}, 1)$, we have $n^{(\sqrt{\gamma}+\sqrt{r})^2/\alpha-1/\alpha-1} \to \infty$ as $n\to\infty$. Therefore, we obtain that $A_{n,1}\to\infty$ with probability tending to 1 as $n\to\infty$.

Next consider the part $A_{n,2} = a_{n;\alpha}\left(\sum_{j\in S} w_j - c_0 n^{-1}\right)\min_{i\in S} F^{-1}[1 - p(X_i)]$. Let $\epsilon_n = n^{\gamma_0}$ with $(\gamma - 1/\alpha)\min(\alpha,1) < \gamma_0 < -\gamma$. Parts 2 and 3 of Assumption 4 imply that $\min_{i\in S}|X_i|$ is greater than ϵ_n with probability tending to 1 as $n\to\infty$, because

$$\begin{aligned}\Pr\left(\min_{i\in S}|X_i| < \epsilon_n\right) &\leq \sum_{i\in S}\Pr\left(|X_i| < \epsilon_n\right) = n^\gamma \Pr\left(|X_i| < \epsilon_n\right) \\ &= n^\gamma\left[2\epsilon_n\phi(\mu_0 - \epsilon_n)\right] = n^\gamma\epsilon_n O(e^{-r\log n}) = o(1),\end{aligned} \tag{S1.1}$$

where $\phi(\cdot)$ is the probability density function of a standard normal random variable. Since $F^{-1}\left[1-p(\cdot)|\alpha,\beta\right]$ is increasing, equation (S1.1) is equivalent to the statement that $F^{-1}[1-p(\min_{i\in S}|X_i|)|\alpha,\beta]$ is greater than any $F^{-1}[1-p(\epsilon_n)|\alpha,\beta]$ with probability 1 as $n\to\infty$. Since $\epsilon_n \to 0^+$ as $n\to\infty$, we can apply Lemma 2 to find a lower bound of $F^{-1}[1-p(\epsilon_n)|\alpha,\beta]$, which leads to a lower bound of $|A_{n,2}|$ as follows:

$$\Pr\left\{|A_{n,2}| > a_{n;\alpha}(n^{\gamma-1} - c_0 n^{-1})|\tilde{g}(\epsilon_n)|\right\} \to 1.$$

Note that $e^{\epsilon_n^2/(2\alpha)} \to 1$ as $n \to \infty$. With the assumption that there is a positive constant c_0 such that $\min_{i=1}^n w_i \geq c_0/n$, we have $a_n < n^{1-1/\alpha} c_0^{-1}$, and hence,

$$\begin{aligned} a_n(n^{\gamma-1} - c_0 n^{-1})|\tilde{g}(\varepsilon_n)| &< \tilde{c}_{\alpha,\beta} a_n n^{\gamma-1} \epsilon_n^{-1/\min(\alpha,1)} e^{\epsilon_n^2/(2\alpha)} \\ &\leq \tilde{c}_{\alpha,\beta} c_0^{-1} n^{-1/\alpha+\gamma-\gamma_0/\min(\alpha,1)} e^{\epsilon_n^2/(2\alpha)} \\ &= o(1). \end{aligned}$$

Therefore, $A_{n,2} > o_p(1)$, which completes the proof of the statement that $A_n \to \infty$ with probability 1 as $n \to \infty$.

Next, we show B_n cannot be arbitrary large negative. Under Part 1 of Assumption 4, Theorem 1 implies that as $n \to \infty$,

$$\begin{aligned} \left(\frac{\sum_{j=1}^n w_j^\alpha}{\sum_{k\in S^c} w_k^\alpha}\right)^{1/\alpha} B_n &= \left(\frac{\sum_{j=1}^n w_j^\alpha}{\sum_{k\in S^c} w_k^\alpha}\right)^{1/\alpha} a_{n;\alpha} \sum_{i\in S^c} w_i F^{-1}\left(1 - p_i|\alpha,\beta\right) \\ &\xrightarrow{d} W_{0;\alpha,\beta}, \end{aligned}$$

where $W_{0;\alpha,\beta}$ follows $\boldsymbol{S}(\alpha,\beta)$.

Let $\delta_{\epsilon_n} = \left[\frac{1-\beta}{\pi\epsilon_n}\Gamma(\alpha)\sin\left(\frac{\pi\alpha}{2}\right)\left(\frac{\sum_{k\in S^c} w_k^\alpha}{\sum_{j=1}^n w_j^\alpha}\right)\right]^{1/\alpha}$, where $\epsilon_n = n^{\gamma_0}$ with $(\gamma - 1/\alpha)\min(\alpha,1) < \gamma_0 < -\gamma$. We first consider the $-1 < \beta < 1$ case. Notice that as $n \to \infty$, $\epsilon_n \to 0$ and $\delta_{\epsilon_n}\left(\frac{\sum_{j=1}^n w_j^\alpha}{\sum_{k\in S^c} w_k^\alpha}\right)^{1/\alpha} = \left[\frac{1-\beta}{\pi\epsilon_n}\Gamma(\alpha)\sin\left(\frac{\pi\alpha}{2}\right)\right]^{1/\alpha} \to \infty$. According to the tail approximation of Theorem 1.2 of Nolan (2020), when $0 < \alpha < 2$ and $-1 \le \beta < 1$,

$$\begin{aligned}\Pr\left(B_n < -\delta_{\epsilon_n}\right) &\sim \Pr\left[W_{0;\alpha,\beta} < -\delta_{\epsilon_n}\left(\frac{\sum_{j=1}^{n} w_j^{\alpha}}{\sum_{k\in S^c} w_k^{\alpha}}\right)^{1/\alpha}\right] \\ &\sim \frac{1-\beta}{\pi}\Gamma(\alpha)\sin\left(\frac{\pi\alpha}{2}\right)\delta_{\epsilon_n}^{-\alpha}\left(\frac{\sum_{j=1}^{n} w_j^{\alpha}}{\sum_{k\in S^c} w_k^{\alpha}}\right)^{-1} \\ &= \epsilon_n,\end{aligned} \tag{S1.2}$$

for large enough n. Equation (S1.2) implies that for any $\varepsilon > 0$, there exists a positive constant $\delta > \delta_{\epsilon_n}$ such that $\Pr\left(B_n < -\delta\right) < \varepsilon$ for large enough $n > \varepsilon^{1/\gamma_0}$.

When $\beta = 1$, the distribution is totally skewed to the right, and consequently, for all $i \in S^c$, $\Pr\left(W_{i;\alpha,1} < -x\right) < \Pr\left(W_{i;\alpha,\beta} < -x\right)$ for any $\beta < 1$ and for any large enough x. Therefore, a similar argument to the one for $-1 \le \beta < 1$ holds for $\beta = 1$. That is, B_n cannot be arbitrary large negative for all $0 < \alpha < 2$ and $-1 < \beta \le 1$, which finishes the proof. □

S2 Simulation Results

In this section, we present the setting and results for our simulations. Similarly to Liu and Xie (2020), a collection of test scores, $\boldsymbol{X}$, is drawn from $N_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. All diagonal elements of the covariance matrix $\boldsymbol{\Sigma}$ are set as one. There are four models for the covariance matrix $\boldsymbol{\Sigma}$ considered to represent different dependence structures. Model 1 is the scheme where the individual tests are independent. In Models 2, 3, and 4, the off-diagonal entries of the covariance matrix $\mathbf{\Sigma} = (\sigma_{ij})$ are functions of ρ.

1. Independent. The correlation between each pair of underlying test scores is zero, i.e., $\mathbf{\Sigma} = I_n$.

2. AR(1) correlation. The correlation between a pair of underlying test scores decays exponentially fast as their distances increase; $\sigma_{ij} = \rho^{|i-j|}$.

3. Exchangeable structure. The correlation between each pair of underlying test scores $\sigma_{ij} = \rho$ for all $i \neq j$.

4. Polynomial decay. The correlation between the ith and jth test scores, σ_{ij}, is set to be $\frac{1}{0.7+|i-j|^{\rho}}$. It should be noted that the correlation is a decreasing function of ρ, unlike Models 2 and 3 above.

The simulation is conducted in `R`. `qstable` function in `stabledist` package (Wuertz et al., 2016) is used to calculate quantiles of stable distributions. We truncate too small and too large p-values at 10^{-6} and $1-10^{-6}$, respectively. This is to avoid technical issues involved with too large quantiles in absolute values in the `qstable` function. The number of Monte Carlo replications is 1000. The number of individual tests in each Monte Carlo replication is 40 ($n = 40$). The significance level is set to be 5%. The parameter ρ that governs the strength of the dependencies is set to be $0.2, 0.4, 0.6$, or 0.8. Note that larger ρ implies stronger dependencies in Models 2 and 3, and weaker dependencies in Model 4. For the SCT, all combinations of $\alpha = 0.1, 0.3, \dots, 1.9$ and $\beta = -0.8, -0.6, \dots, 1$ are considered in addition to $(\alpha, \beta) = (1, 0)$, which is equivalent to the CCT. We also consider Stouffer's Z-score, which would correspond to the SCT with $\alpha = 2$ and $\beta = 0$. Note that although Stouffer's Z-score can be written in the SCT form, Stouffer's Z-score is not a part of the SCT family we consider in our paper. The test statistics are calculated as equation (0.1) with equal weights $w_i = 1/n$.

When calculating the sizes, data are generated from a multivariate normal distribution with mean $\boldsymbol{\mu} = \mathbf{0}$. For powers, following one of the sparse alternative setting in Liu and Xie (2020), randomly chosen 4 indices in each replication are set to have mean 2.095662. This choice corresponds to $\gamma = \log(4)/\log(40) \approx 0.3758$ and $r = 2.095662/\sqrt{2\log(40)} \approx 0.5953$ so that $\mu_i = \sqrt{2r\log n} = 2.095662$ and $[n^\gamma] = 4$. Note that $(\sqrt{r} + \sqrt{\gamma})^2 \approx 1.917 > \max(\alpha, 1)$ for all αs considered in our simulation, satisfying Part 3 of Assumption 4. However, Part 2 of Assumption 4 does not hold for $\alpha > 1/(2\gamma) \approx 1.33$. For raw powers, the cutoff values are taken directly from the corresponding stable distributions. For the size-adjusted powers, 1000 Monte Carol replications are first drawn under the global null

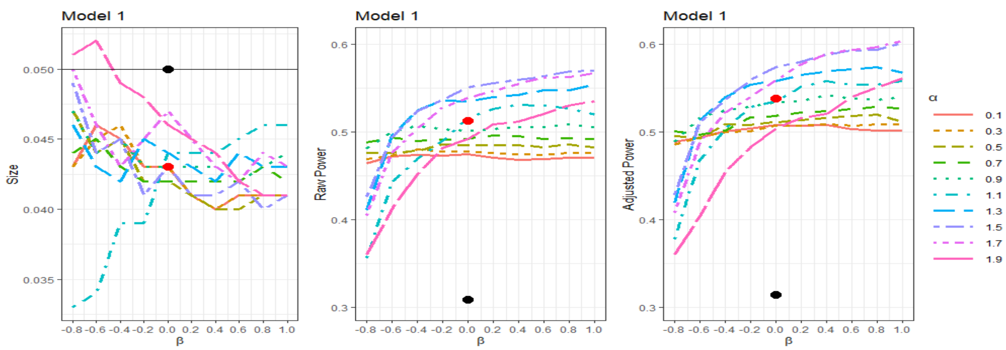

Figure 1: Sizes, raw powers and size-adjusted powers of Model 1 where tests are independent. Lines indicate the SCT with different αs and βs. Red and black dots represent the CCT (SCT with $\alpha = 1$ and $\beta = 0$) and Stouffer's Z-score, respectively.

hypothesis. Combined test statistics are calculated for each Monte Carlo replication. The simulation-based cutoff for each method is determined as the 95% quantile of the 1000 test statistics. After that, another set of 1000 Monte Carlo replications is drawn under the sparse alternative. The proportion of test statistics that are greater than the simulation-based cutoffs is the size-adjusted powers.

Figures 1-4 present the sizes, raw powers, and size-adjusted powers for different αs and βs under the four models. Red dots indicate the CCT, which corresponds to the SCT with $\alpha = 1$ and $\beta = 0$. Black dots indicate Stouffer's Z-scores. The black solid lines in the size plots represent the nominal significance level 0.05.

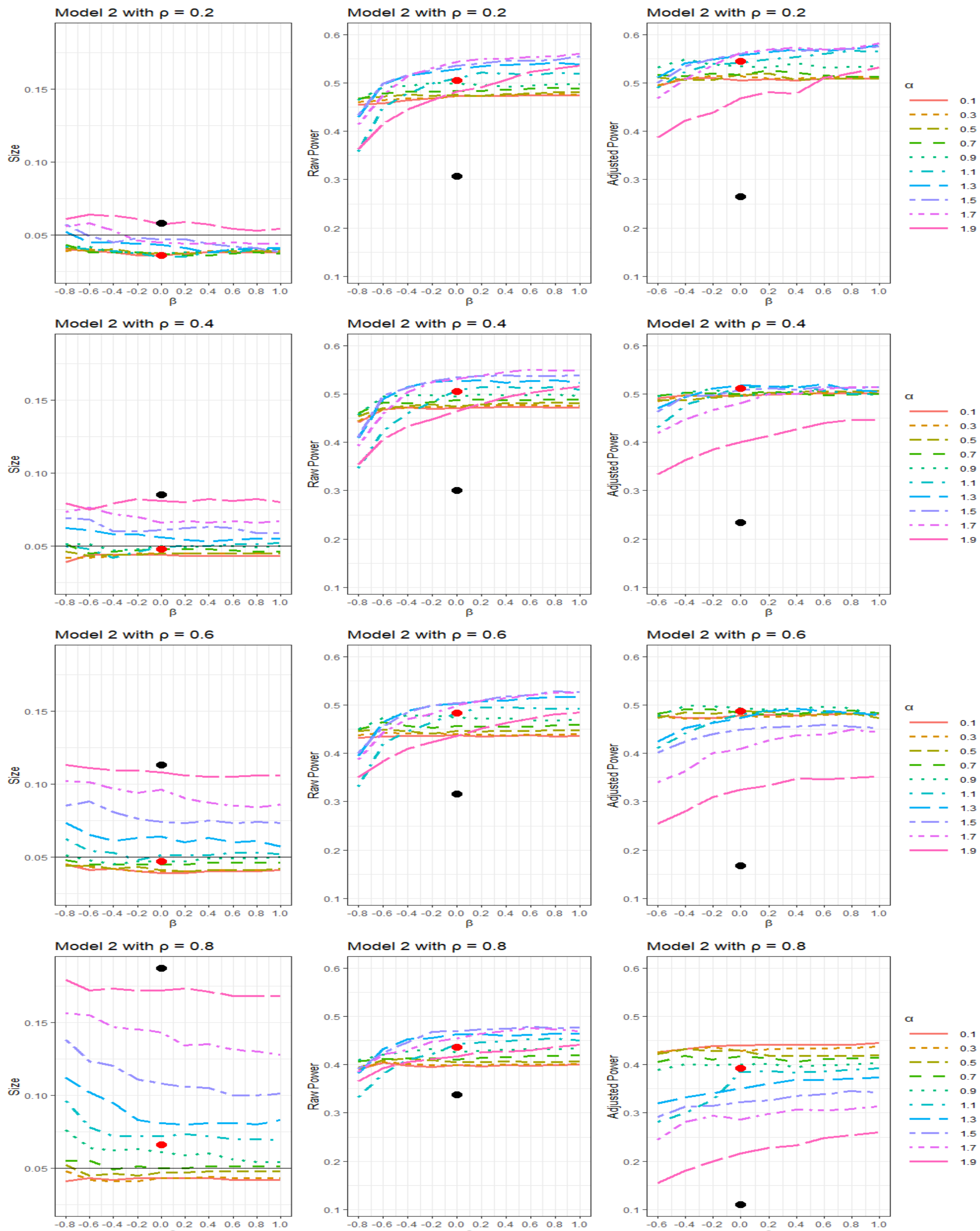

Figure 2: Sizes, raw powers and size-adjusted powers of Model 2 where tests are correlated with AR(1) correlation structures with different ρs. Lines indicate the SCT with different αs and βs. Red and black dots represent the CCT (SCT with $\alpha = 1$ and $\beta = 0$) and Stouffer's Z-score, respectively.

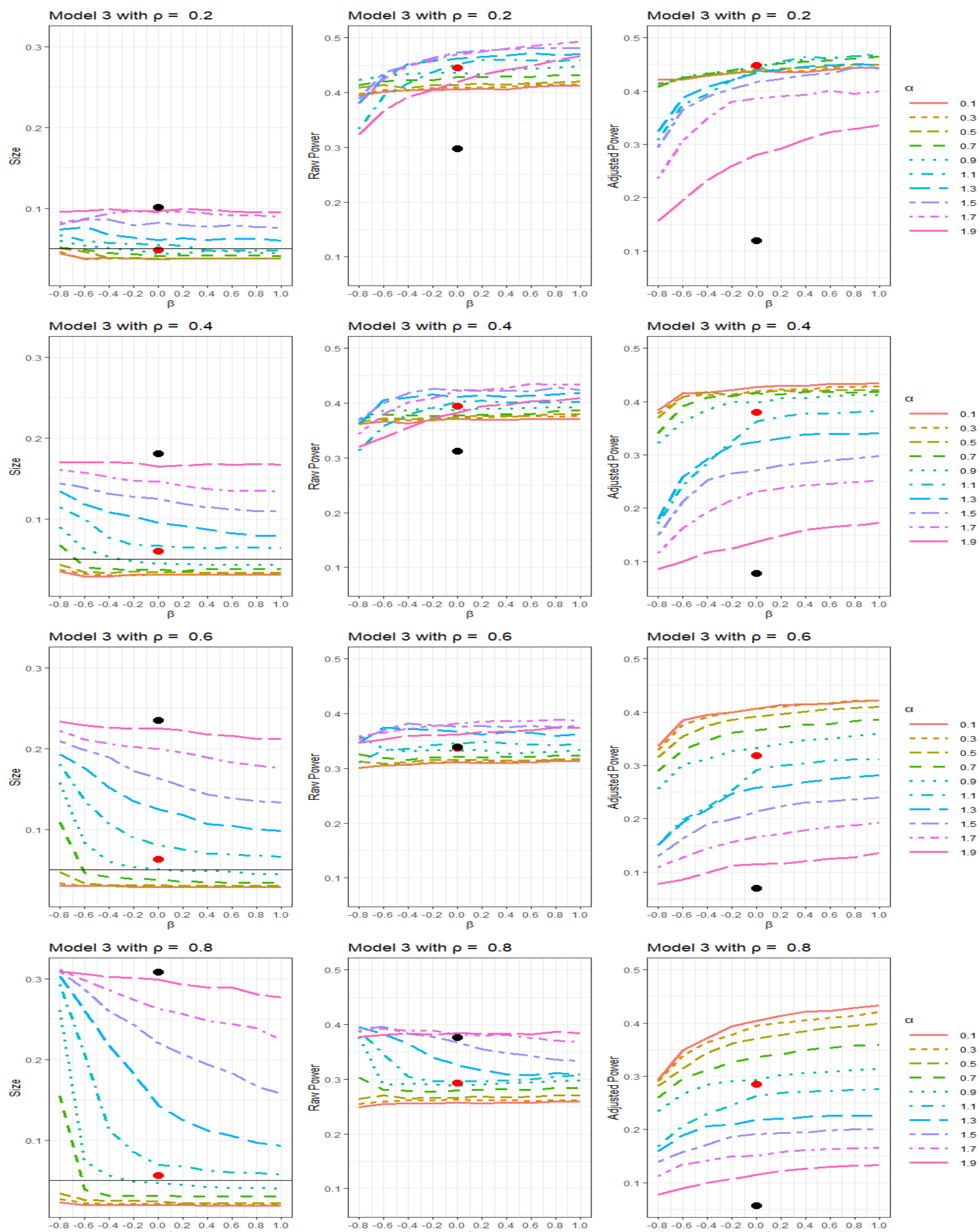

Figure 3: Sizes, raw powers and size-adjusted powers of Model 3 where tests are correlated with exchangeable correlation structures with different ρs. Lines indicate the SCT with different αs and βs. Red and black dots represent the CCT (SCT with $\alpha = 1$ and $\beta = 0$) and Stouffer's Z-score, respectively.

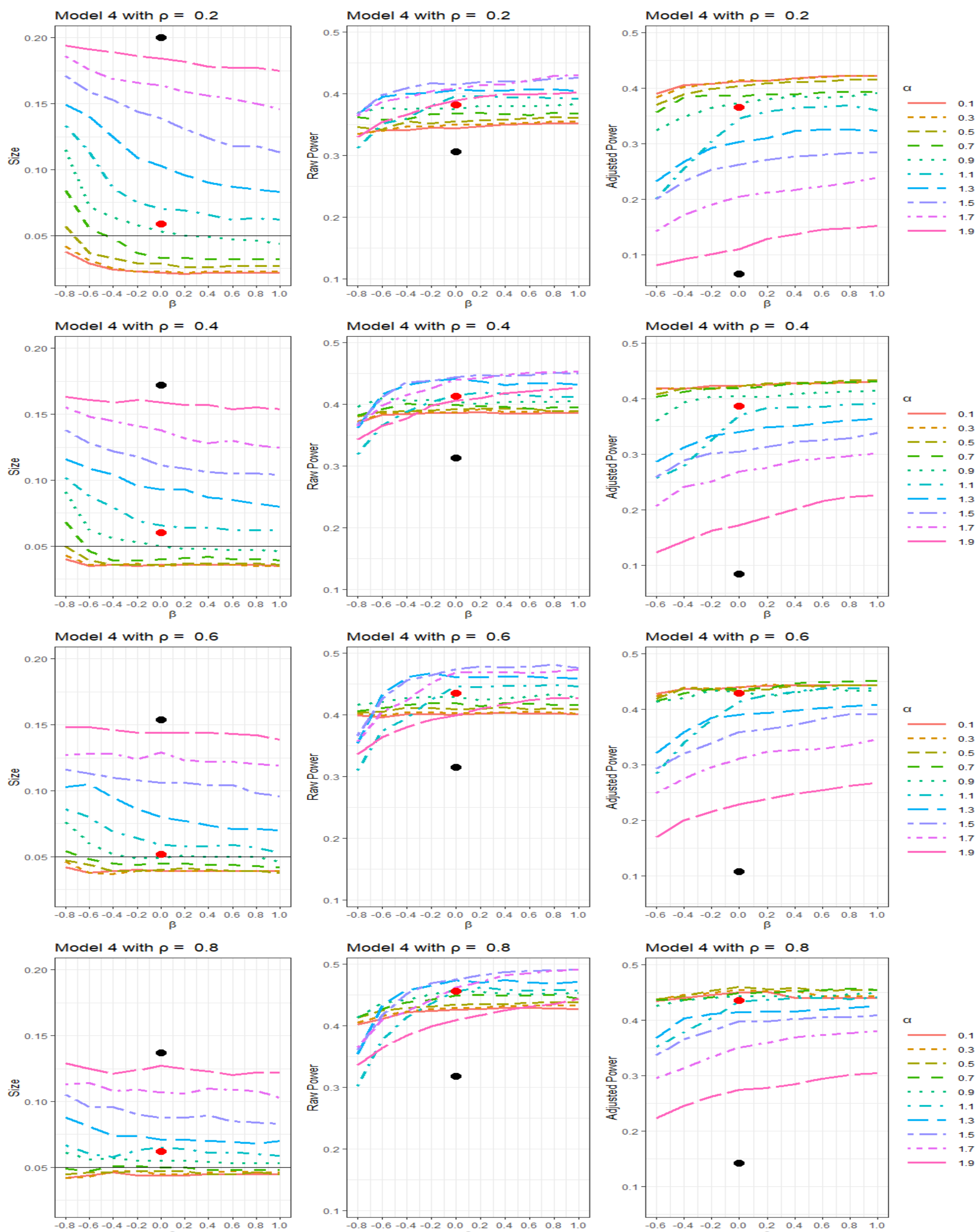

Figure 4: Sizes, raw powers and size-adjusted powers of Model 4 where tests are correlated as polynomial decayed correlation structures with different ρs. Lines indicate the SCT with different αs and βs. Red and black dots represent the CCT (SCT with $\alpha = 1$ and $\beta = 0$) and Stouffer's Z-score, respectively.

Bibliography

Cai, T. T., W. Liu, and Y. Xia (2014). Two-sample test of high dimensional means under dependence. *Journal of the Royal Statistical Society: Series B: Statistical Methodology 76*(2), 349–372.

Liu, Y. and J. Xie (2020). Cauchy combination test: a powerful test with analytic p-value calculation under arbitrary dependency structures. *Journal of the American Statistical Association 115*(529), 393–402.

Nolan, J. P. (2020). *Univariate stable distributions.* Springer.

Wuertz, D., M. Maechler, and R. core team members. (2016). *stabledist: Stable Distribution Functions.* R package version 0.7-1.