

Supplementary Materials to “Large-Scale Hypothesis Testing for Causal Mediation Effects with Applications in Genome-wide Epigenetic Studies”

Proof of the Three Results

Proof of Result 1

First, it can be easily shown that

$$T_{Sobel}^2 = \frac{Z_\beta^2 Z_\gamma^2}{Z_\beta^2 + Z_\gamma^2} = \left(Z_\beta^{-2} + Z_\gamma^{-2} \right)^{-1}.$$

In Case 3, both Z_β and Z_γ follow $N(0, 1)$, so T_{Sobel}^2 follows $\frac{1}{4}\chi_1^2$ using Corollary 1 of Girón and del Castillo (2001). There is another proof using the “super Cauchy phenomenon” (Pillai and Meng 2016). Note that Z_β^{-2} and Z_γ^{-2} are independent standard Lévy random variables with characteristic functions $\exp(-\sqrt{-2it})$, where $\mathbf{i}^2 = -1$. Then we have $Z_\beta^{-2} + Z_\gamma^{-2} \sim 4\chi_1^{-2}$. This completes the proof.

The power of the Sobel’s test can be easily obtained by using the definition of power function under the alternative hypothesis and can be written as a two-fold integral. Using the Cauchy-Schwarz inequality for integrals (Steele 2004, p.11), its power is maximized when $|u_\beta(n)| = |u_\gamma(n)|$.

Proof of Result 2

We are interested in testing $H_0 : \beta\gamma = 0$ versus $H_a : \beta\gamma \neq 0$. Denote the null parameter space as Θ_0 and the full parameter space as Θ . The likelihood function $L(\theta|Y, M, A, X)$ for the outcome and mediator regressions can be written as

$$L(\beta_0, \beta_A, \beta, \beta_X^T, \sigma_Y^2, \gamma_0, \gamma, \gamma_X^T, \sigma_M^2 | Y, M, A, X) = L(\beta_0, \beta_A, \beta, \beta_X^T, \sigma_Y^2 | Y, M, A, X) L(\gamma_0, \gamma, \gamma_X^T, \sigma_M^2 | M, A, X).$$

The likelihood ratio test is

$$\lambda = 2 \log \frac{\sup_{\Theta} L(\theta | Y, M, A, X)}{\sup_{\Theta_0} L(\theta | Y, M, A, X)}.$$

The parameters in the outcome and mediator regression models can be estimated using maximum likelihood estimators (MLE). Denote $L_Y = L(\beta_0, \beta_A, \beta, \beta_X^T, \sigma_Y^2 | Y, M, A, X)$ which corresponds to the outcome regression and $L_M = L(\gamma_0, \gamma, \gamma_X^T, \sigma_M^2 | M, A, X)$ which corresponds to the mediator regression.

When $\beta = 0$,

$$\lambda = 2 \log \frac{\sup_{\Theta} L_Y L_M}{\sup_{\Theta_0} L_Y L_M} = 2 \log \frac{\sup_{\Theta} L_Y}{\sup_{\Theta_0} L_Y} = \lambda_Y \rightarrow \chi_1^2,$$

where we only test whether the parameter β is zero or not using the outcome regression model and the parameter γ is unrestricted.

When $\gamma = 0$,

$$\lambda = 2 \log \frac{\sup_{\Theta} L_Y L_M}{\sup_{\Theta_0} L_Y L_M} = 2 \log \frac{\sup_{\Theta} L_M}{\sup_{\Theta_0} L_M} = \lambda_M \rightarrow \chi_1^2,$$

where we only test whether the parameter γ is zero or not using the mediator regression model and the parameter β is unrestricted. Because the null likelihood is in the denominator, so the LRT is $\lambda = \min(\lambda_Y, \lambda_M)$ and we reject the null when $\min(\lambda_Y, \lambda_M) > C_\alpha$, where C_α is the critical value. Define $p_\beta = Pr(\lambda_Y > \lambda_Y^{obs})$ and $p_\gamma = Pr(\lambda_M > \lambda_M^{obs})$, then the MaxP rejects the null if $\max(p_\beta, p_\gamma) < \alpha$.

$Power = P(MaxP < \alpha; H_a) = Pr(p_\beta < \alpha; H_a) Pr(p_\gamma < \alpha; H_a)$. Using inequality $ab \leq \frac{a^2+b^2}{2}$, we know the power of MaxP is maximized when the powers for testing β and γ are equal, that is when $|u_\beta(n)| = |u_\gamma(n)|$.

Proof of Result 3

In the null Case 1, since $\hat{w}_j, j = 1, 2, 3$ are consistent, then $\hat{w}_1 \rightarrow w_1$, and \hat{w}_2, \hat{w}_3 converge to zero in probability. Then DACT is equivalent to p_γ asymptotically when m goes to infinity. Combined this with the analysis results in Section 4, DACT is equivalent to the Sobel's test and the MaxP test asymptotically in the null Case 1. The same argument applies to the null Case 2.

In the null Case 3, again because $\hat{w}_j, j = 1, 2, 3$ are consistent, the proposed DACT is equivalent to $\{MaxP\}^2$. Note that $Pr(\{MaxP\}^2 < \alpha) = Pr(MaxP < \sqrt{\alpha}) = \alpha$ because MaxP follows the $Beta(2, 1)$ distribution. Hence, the proposed DACT has the correct size in the null Case 3. The actual size of the Sobel's test is $2\Phi(2z_{\alpha/2}) < 2\Phi(z_{\alpha/2}) = \alpha$, where $z_{\alpha/2}$ denotes the lower $\alpha/2$ quantile of a standard normal distribution. The actual size of the MaxP test is $Pr(MaxP \leq \alpha) = \alpha^2 < \alpha$ for $\alpha \in (0, 1)$.

We now prove that the multiple testing procedure in Result 3(c) can control the local FDR. The marginal density f can be consistently estimated using kernel density estimator (Wasserman

2006, pp. 133). Because $\widehat{\pi}_0^{DACT}$, \widehat{f}_0 , \widehat{f} are consistent estimates of π_0^{DACT} , f_0 , f respectively, hence the local FDR $\widehat{fdr}(z) = \widehat{\pi}_0^{DACT} \widehat{f}_0(z) / \widehat{f}(z)$ is also consistent using the continuous mapping theorem (van der Vaart 2000, pp. 7). Denote \widehat{z}_q as the estimated threshold, then we have \widehat{z}_q converges to z_q in probability using Lemma A.5 of Sun and Cai (2007). Then we have $\widehat{fdr}(\widehat{z}_q) \rightarrow fdr(z_q) \leq q$ in probability. For the tail area FDR, we can simply replace \widehat{f}_0 , \widehat{f} by \widehat{F}_0 , \widehat{F} respectively in the above proof. This completes the proof.

References

- Girón, F. J. and del Castillo, C. (2001). A note on the convolution of inverted-gamma distributions with applications to the behrens-fisher distribution. *RACSAM* **95**, 39–44.
- Pillai, N. S. and Meng, X.-L. (2016). An unexpected encounter with Cauchy and Lévy. *The Annals of Statistics* **44**, 2089–2097.
- Steele, J. M. (2004). *The Cauchy-Schwarz master class : an introduction to the art of mathematical inequalities*. MAA problem books series. Cambridge University Press, Cambridge ; New York.
- Sun, W. and Cai, T. T. (2007). Oracle and adaptive compound decision rules for false discovery rate control. *Journal of the American Statistical Association* **102**, 901–912.
- van der Vaart, A. (2000). *Asymptotic Statistics*. Asymptotic Statistics. Cambridge University Press.
- Wasserman, L. (2006). *All of Nonparametric Statistics*. Springer Texts in Statistics. Springer New York.