Received: 14 Sep 2024 | Revised: 11 Dec 2025 | Accepted: 19 Dec 2025 | Available online: 21 May 2026
https://doi.org/10.13057/ijas.v8i2.93376

E-ISSN: 2621-086X
https://jurnal.uns.ac.id/ijas

Indonesian Journal of Applied Statistics Vol. 8, No.2, pp. 175-181, 2025

Exploring Statistical Power and Mediation Analysis: Understanding the Impact of Antecedent-Mediator-Outcome Relationships

Szilárd Nemes*
Institute of Clinical Sciences, Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden.
*Corresponding author: szilard.nemes@outlook.com

Abstract

This paper explores the phenomenon of statistical power stagnation and decline in mediation analysis, specifically focusing on the interplay between the antecedent variable, mediator, and outcome. Mediation analysis is a critical statistical tool used to understand the causal pathways between variables. However, statistical power may not always increase with stronger relationships between the antecedent and mediator, often stagnating or even declining due to variance inflation caused by multicollinearity. We provide a in detail examination of this issue, including key theoretical concepts, the mathematical foundations of variance inflation, and the impact of mediator-antecedent correlations on power. A simulation study further illustrates how varying these correlations affects statistical power, variance estimates, and possible bias in mediation effects. Our findings indicate that while increasing the strength of the relationship between the antecedent and mediator improves mediation detection initially, beyond a certain threshold, it results in inflated variance estimates, leading to decreased precision and power. Variance inflation of the mediated effect is more accentuated than variance inflation of regression coefficients.

Keywords: mediation; variance inflation; Sobel test

How to Cite:
S. Nemes, “Exploring statistical power and mediation analysis: understanding the impact of antecedent-mediator-outcome relationships,” *Indonesian Journal of Applied Statistics,* vol. 8, no. 2, pp. 175-181, 2025, doi: 10.13057/ijas.v8i2.93376.

1. INTRODUCTION

Mediation analysis seeks to elucidate the underlying structural relationships between an antecedent variable and an outcome by incorporating a third variable, known as the mediator. The mediator serves to clarify the nature of the relationship between the antecedent and the outcome. Mediation is established when the antecedent influences the mediator, and a subsequent change in the mediator is associated with a change in the outcome, contingent on the presence of the antecedent variable [1],[2]. The mediator thus partially or fully accounts for the relationship between the antecedent and the outcome, allowing the total effect of the antecedent on the outcome to be decomposed into mediated and non-mediated paths [3],[4].

In mediation studies, the focus is primarily on the mediated path, with researchers often employing null-hypothesis tests to assess its significance [5],[6]. Regardless of the testing methodology used, the null hypothesis typically posits that the effect size of the mediated path is zero. A wide range of significance tests has been developed and applied to evaluate mediation effects. Numerical and simulation studies have revealed notable insights, particularly that the statistical power of mediation tests may asymptote below 1 and, in some cases, may even decline. Fritz et al. [5] found that this decline in power primarily occurs when the relationship between the mediator and the outcome is weak, while the relationship between the antecedent variable and the mediator is strong. This induces a multicollinearity in the regression equation which leads inflated variance estimates and decrease in statistical power [7].

In this note, we explore the causes of power stagnation and decline. We start by introducing key concepts and definitions, followed by a discussion of the theoretical foundations underlying power stagnation. Building on Beasley's [7] work, we present closed-form solutions and further investigate these findings through a simulation study. We conclude with a brief discussion of the results.

2. METHODS

We commence with the definition of a single mediator model when the effect of the antecedent variable X on the outcome Y is mediated my M. Mathematically, this model is represented by three regression equations

$$Y = \alpha_1 + cX + \varepsilon_1 \quad (1)$$

$$M = \alpha_2 + aX + \varepsilon_2 \quad (2)$$

$$Y = \alpha_3 + c^{'}X + bM + \varepsilon_3 \quad (3)$$

where c is the total effect of the antecedent variable on the outcome, a is the effect of the antecedent variable on the mediator while b is effect of the mediator on the outcome adjusted for the effect of the antecedent variable. The casual steps strategy advocated by Baron and Kenny [8] uses estimates form all three models to test for mediation effect; however, this approach has low statistical power. Alternatively, product of coefficient tests such as Sobel-tests [9],[10] can be used. The product of coefficient tests proceeds with simplification of equations (2) and (3) by replacing M in equation (3) with the structural component from equation (2) leading to

$$Y = (\alpha_3 + b\alpha_2) + X\left(c^{'} + ab\right) + \varepsilon_3 \quad (4)$$

The limiting distribution of the effect size estimator for the mediation effect ab is obtained by first order Delta-method, giving $\hat{a}\hat{b} \sim N\left(ab, b^2\sigma_a^2 + a^2\sigma_b^2\right)$ and

$$Z = \frac{ab}{\sqrt{b^2\sigma_a^2 + a^2\sigma_b^2}} \quad (5)$$

where Z has standard normal distribution with expectation 0 and variance 1.

The null-hypothesis of the test states that $H_0: ab = 0$ the alternative $H_1: ab \neq 0$.. The rejection region of the test is $|ab| > z\left(\frac{\alpha}{2}\right)\sigma_{ab}$ and the probability of rejecting the null-hypothesis when is false, statistical power is given by

$$1 - \Phi\left\{z_{\frac{\alpha}{2}} - \frac{ab}{\sigma_{ab}}\right\} + \Phi\left\{-z_{\frac{\alpha}{2}} - \frac{ab}{\sigma_{ab}}\right\} \quad (6)$$

where $z_{\alpha/2}$is the quantile of the standard normal distribution at the desired significance level, α.

3. RESULTS AND DISCUSSION

3.1 Factors Influencing Statistical Power in Mediation Analysis

Statistical power varies with effect size and sample size. As the sample size increases the precision of regression coefficient increases and their variance decrees. As a result, the signal-to-noise ratio decreases and detection of a mediation effect become easier. Independently of the magnitudes of regression coefficients, statistical power will depend on the ratio between signal $\hat{a}\hat{b}$ and the noise, $\hat{\sigma}_{ab}$. The variance of the mediation, $\hat{\sigma}_{ab}^2$, incorporates the variability both from the antecedent and mediator, $b^2\sigma_a^2 + a^2\sigma_b^2$. The effect of the antecedent variable on the mediator, a, and its variance (σ_a^2) is estimated from equation 2 and their value depends on the relationship between the mediator and the antecedent

https://doi.org/10.13057/ijas.v8i2.93376
E-ISSN: 2621-086X

variable. The stronger the correlation between the two the lower the variance of the regression coefficient will be. This relationship will directly influence $\sigma_b^2$and $\hat{\sigma}_{ab}^2$. For the simple linear regression equation, the regression coefficients can be estimated by minimizing the residual sum of squares

$$argmax\left\{\sum_i \left(y_i - \alpha_3 - c'x_i - bm_i\right)^2\right\} \tag{7}$$

or by solving the normal equations, $(\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{y}$ where $\boldsymbol{X}$ is the design matrix compromising a vector of ones for the intercept a vector with the measurements for the antecedent variable and a vector with the measurements for the mediator. The variance of regression coefficients is estimated as $\sigma^2(\mathbf{X}^T\mathbf{X})^{-1}$ with

$$\sigma^2 = \frac{\sum_i \left(y_i - \alpha_3 - c'x_i - bm_i\right)^2}{n-3} \tag{8}$$

If the regression equation contains an antecedent variable and a mediator that are highly correlated to each other then $\boldsymbol{X}^T\boldsymbol{X}$ will have a determinant close to zero, making inversion imprecise and inflates the variance of the parameter estimates. This variance inflation depends on the strength of the relationship between the antecedent variable and mediator and is estimated by mediator. The estimated variance for the regression coefficient b in practice will be

$$\sigma_b^2 = \frac{\sigma^2}{(1 - R_{m,x}^2)\sum_i (m_i - \bar{m})^2} \tag{9}$$

Here, h $R_{m.x}^2$ is the squared correlation coefficient between the antecedent and mediator variable and $(1 - R_{m,x}^2)^{-1}$ is the variance inflation factor, that measures the increase in variance due to collinearity [11]. Additionally with some algebra it can be shown that σ_a^2equals

$$\frac{\left(1 - R_{m,x}^2\right)}{n-2}\frac{\sigma_m^2}{\sigma_x^2} \tag{10}$$

Is easy to see if $R_{m.x}^2 = 0$ (i.e. no relationship between the mediator and antecedent variable) then σ_b^2 will depend solely on the relationship between mediator and outcome. On the other extreme if $R_{m.x}^2 = 1$ then $\sigma_b^2$is not defined and $\lim_{R_{m.x}^2 \to 1} \sigma_{ab}^2 = \infty$. This because the mediator and antecedent variable are linear combinations of each other and contain exactly the same information about the outcome. In practical applications $R_{m.x}^2$ will be somewhere between 0 and 1. The value of $R_{m.x}^2$will give insight in how much information about the outcome is shared between the antecedent variable and mediator. This shared information contributes to the inflation of the variance estimators for the coefficients in equation (3) as the regression equation cannot attribute this information to either the antecedent variable or the mediator thus increasing the insecurity in the estimation. The more information the two share higher the inflation variance.

In a mediation context, the strength of the relationship between the antecedent variable and the mediator plays a critical role, simultaneously enhancing the mediation effect while also introducing potential complexities. By definition we wish for a strong antecedent-mediator relationship, but at the same time the stronger the relationship gets the larger the variance inflation will be. Consequently, there is an inverse relationship between σ_b^2 and σ_a^2. If one increases the other inevitably will decrease. If the antecedent does not affect the mediator (e.g. no mediation effect) then σ_b^2 depends only on the strength of association between the alleged mediator and outcome. If there is a mediation effect, then σ_b^2 will be inflated by a factor of $(1 - R_{m,x}^2)^{-1}$. As $R_{m.x}^2$ increases so does σ_a^2 decrease and σ_b^2 get more and more inflated. Ultimately the achieved power depends on three parameters, the correlation between the mediator and antecedent ($r_{m,x}$), correlation between the mediator and outcome ($r_{m,y}$) and the correlation between the antecedent and outcome ($r_{x,y}$).

https://doi.org/10.13057/ijas.v8i2.93376
E-ISSN: 2621-086X

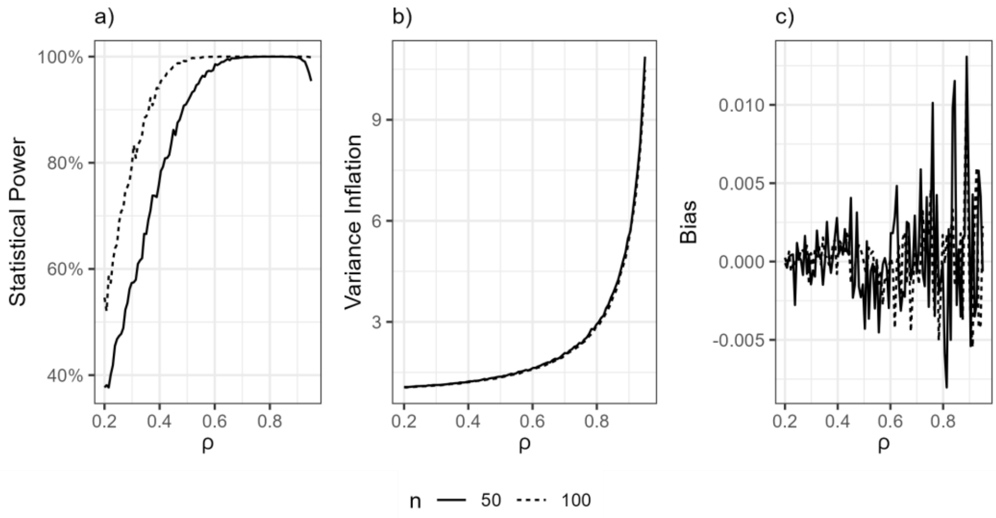

Figure 1. The effect of the correlation between the antecedent and mediator variables on the statistical power for detecting mediation (a), variance inflation (b) and the bias of the estimated mediated effect.

Variance inflation in the b-coefficient propagates directly into the variance of the mediated effect. Similar phenomena have been demonstrated in modern psychometric mediation work, where imprecise mediator measurement inflates standard errors and reduces power

3.2 Numerical Evaluations

Similar to Zhang [12] whi has shown that product-based mediation tests suffer from structural conservatism and variance irregularities we examine the finite-sample operating characteristics of mediation estimators under structural configurations known to distort Type I error and reduce power. This section provides a numerical evaluation of the results detailed in the previous section. The simulation was conducted following the established guidelines outlined by Burton et al [13]. For this analysis, we considered antecedent variables that follow a standard normal distribution, and a zero mean mediator variable with the correlation coefficient between (ρ) them varying incrementally from 0.2 to 0.95 in steps of 0.0075. Using the relationship between correlation and regression coefficients, the variance of the mediator variable was set to $\sqrt{\rho}^{-1}$, thus $a = 1$. The outcome variable was generated in accordance with Equation 3, using parameters $c' = 0.2$ and $b = 1$ and mean zero residuals with standard deviation of 0.5. We analyzed two different sample sizes: 50 and 100. For each combination of sample size and correlation coefficient, 1000 data sets were generated. Linear regression was applied to each data set to estimate the coefficients from Equations 1, 2, and 3, as well as their associated variances. These estimates were then utilized as inputs in Equation 6. Additionally, we assessed variance inflation as a function of the correlation between the antecedent and mediator variables. Finally, we investigated the potential impact of multicollinearity and variance inflation on the bias of the estimated mediated effect. As this is a simulation study with a known population value, the bias is given by $bias(\hat{a}\hat{b}) = E[\hat{a}\hat{b}] - ab$ and its associated variance is σ^2_{ab}. Figure 1 summarizes the results of the simulation. As expected, statistical power initially increases with increasing correlation between antecedent and mediator, however after a peak is passed around $\hat{\rho} \approx 0.6$, statistical power decreases. This coincides with a variance inflation factor of around 1.6, which suggests that the standard error for $\hat{b}$ is around 26% larger than it would have been without collinearity between antecedent and mediator. Furthermore, the results of the simulations suggest that the accuracy of the estimates is not affected by the correlation between antecedent and mediator (Figure 1 c).

https://doi.org/10.13057/ijas.v8i2.93376
E-ISSN: 2621-086X

Table 1 provides additional insights into how the correlation between antecedent and mediator variables induces variance inflation. As expected, when the mediator is regressed on the antecedent variable, a stronger correlation between the two results in a lower variance error for the regression coefficient ($\hat{\sigma}_a^2$). This indicates increased precision in estimating the effect of the antecedent on the mediator. However, when the outcome variable is regressed on both the mediator and the antecedent, the precision of the estimate for $\hat{b}$ significantly decreases. This reduced precision is then carried over to the estimates of the mediation effect ($\hat{a}\hat{b}$). Since the variance estimator for the mediated effect incorporates both $\hat{a}$ and $\hat{b}$ as multiplicative factors, the imprecision in estimating the mediated effect exceeds the imprecision in estimating $\hat{b}$.

Table 1. Precision of the variance estimates for regression coefficient and mediated effect as a function of the correlation between antecedent and mediator variables.

VIF	$\widehat{VIF}$	$\hat{\rho}$	$\hat{\sigma}_{ab}^2$	$\hat{\sigma}_a^2$	$\hat{\sigma}_b^2$
1	1.01	0.0209	21.400	21.4000	0.000005
2	2.09	0.7030	0.0269	0.0213	0.00542
5	5.30	0.8901	0.0271	0.0053	0.02170
50	51.2	0.9900	0.2660	0.0004	0.26501

In this note, we examined how power stagnation is a predictable phenomenon, directly resulting from the strength of the relationship between the mediator and the antecedent variable. This phenomenon is not exclusive to mediation analysis but is also observed in any multivariate regression context. Fritz et al. [5] recommended that, when necessary, researchers should prioritize mediators that exhibit the strongest relationship to the outcome over those related to the antecedent. While this is a reasonable suggestion, it may not always be feasible to isolate a single factor in practice. Additionally, as sample size increases or the signal-to-noise ratio decreases, statistical power is expected to asymptotically approach one. This is due to the variance of the mediation effect—essentially the noise that obscures the signal—becoming increasingly negligible. The strength of association, quantified by the correlation between the antecedent and mediator, directly impacts the variance of the mediation effect, but this relationship does not exhibit asymptotic behavior. Recent econometric work also highlights how indirect effects depend critically on structural identification and the interplay between pathways [14]. Our findings complement these results by demonstrating how collinearity specifically inflates the variance of the mediated effect and contributes to power stagnation.

Addressing power stagnation might be challenging. While several alternatives to Sobel's test exist, resampling methods and the distribution of product confidence limits have proven most effective [15] unless computational burden in big data hinders it [16]. In this note, we utilized Sobel's test for illustrative and pedagogical purposes, though we believe the concepts discussed are broadly applicable, regardless of the testing method employed. Alternatives to Sobel's test do not mitigate variance inflation, which is an inherent characteristic of multivariate regression models. Variance inflation arising from strong correlations between predictors has long been recognized to inflate standard errors and obscure true effects [17] in addition there might be multiple correlated mediators within the same study [18]. Naturally, increasing sample size leads to greater statistical power, and in most scenarios, statistical power will asymptotically approach one as the sample size becomes sufficiently large. In practical applications, it's uncommon to have variables that capture different parameters of interest for the research question and exhibit correlations strong enough to raise concerns about statistical power in mediation analysis. Integrative analysis of omics data with mediation has highly correlated mediator and the antecedent variable, however mediation in this field is limited [19],[20].

Mediation analysis is particularly common in social science research, where relationships between variables are often indirect rather than direct. It allows researchers to model these complex relationships by identifying the processes or mechanisms through which one variable influences another, offering

https://doi.org/10.13057/ijas.v8i2.93376
E-ISSN: 2621-086X

valuable insights into the 'how' or 'why' behind observed effects. However, in the social sciences, correlations between variables are typically moderate to weak compared to the stronger correlations often seen in physical sciences. This is due to the multifaceted nature of social phenomena, which are influenced by numerous factors, making it challenging for any single variable to explain a large portion of the variance in outcomes.

4. CONCLUSIONS

In summary, power stagnation and decline in mediation analysis can arise naturally when the antecedent and mediator become highly correlated. Although a stronger antecedent--mediator association increases the mediated effect, it also inflates the variance of the mediator coefficient in the outcome model, reducing precision and eventually statistical power. The numerical results illustrate that this variance inflation affects the mediated effect more strongly than the individual regression coefficients, while not materially biasing the point estimate. Thus, mediation analyses should consider not only the expected size of the indirect effect, but also the loss of precision induced by collinearity between the antecedent and mediator.

REFERENCES

[1] T. J. VanderWeele, "Mediation analysis: a practitioner's guide," *Annual review of public health,* vol. 37, no. 1, pp. 17-32, 2016, https://www.annualreviews.org/content/journals/10.1146/annurev-publhealth-032315-021402.

[2] T. J. VanderWeele, "Subtleties of explanatory language: what is meant by “mediation”?," *European journal of epidemiology,* vol. 26, no. 5, pp. 343-346, 2011, https://doi.org/10.1007/s10654-011-9588-z.

[3] E. Suzuki, E. Yamamoto, and T. Tsuda, "Identification of operating mediation and mechanism in the sufficient-component cause framework," *European journal of epidemiology,* vol. 26, no. 5, pp. 347-357, 2011, https://doi.org/10.1007/s10654-011-9568-3.

[4] H. Lee, R. D. Herbert, and J. H. McAuley, "Mediation analysis," *Jama,* vol. 321, no. 7, pp. 697-698, 2019, https://doi.org/10.1001/jama.2018.21973.

[5] M. S. Fritz, A. B. Taylor, and D. P. Mackinnon, "Explanation of Two Anomalous Results in Statistical Mediation Analysis," (in eng), *Multivariate Behav Res,* vol. 47, no. 1, pp. 61-87, 2012, https://doi.org/10.1080/00273171.2012.640596.

[6] K. J. Preacher and A. F. Hayes, "Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models," *Behavior research methods,* vol. 40, no. 3, pp. 879-891, 2008, https://doi.org/10.3758/brm.40.3.879.

[7] T. M. Beasley, "Tests of mediation: Paradoxical decline in statistical power as a function of mediator collinearity," *The Journal of Experimental Education,* vol. 82, no. 3, pp. 283-306, 2014, https://doi.org/10.1080/00220973.2013.813360.

[8] R. M. Baron and D. A. Kenny, "The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations," *Journal of personality and social psychology,* vol. 51, no. 6, p. 1173, 1986, https://doi.org/10.1037/0022-3514.51.6.1173.

[9] M. E. Sobel, "Asymptotic confidence intervals for indirect effects in structural equation models," *Sociological methodology,* vol. 13, pp. 290-312, 1982, https://doi.org/10.2307/270723.

[10] M. E. Sobel, "Some new results on indirect effects and their standard errors in covariance structure models," *Sociological methodology,* vol. 16, pp. 159-186, 1986, https://doi.org/10.2307/270922.

[11] J. Faraway, "Linear models with R. 2oedition," *Florida: Chapman & Hall/CRC,* 2015.

https://doi.org/10.13057/ijas.v8i2.93376
E-ISSN: 2621-086X

[12] H. Zhang, "Efficient adjusted joint significance test and sobel-type confidence interval for mediation effect," *Structural Equation Modeling: A Multidisciplinary Journal,* vol. 32, no. 1, pp. 93-104, 2025, https://doi.org/10.1080/10705511.2024.2392139.

[13] A. Burton, D. G. Altman, P. Royston, and R. L. Holder, "The design of simulation studies in medical statistics," *Statistics in medicine,* vol. 25, no. 24, pp. 4279-4292, 2006, https://doi.org/10.1002/sim.2673.

[14] M. S. Schuler, D. L. Coffman, E. A. Stuart, T. Q. Nguyen, B. Vegetabile, and D. F. McCaffrey, "Practical challenges in mediation analysis: a guide for applied researchers," *Health Services and Outcomes Research Methodology,* vol. 25, no. 1, pp. 57-84, 2025, https://doi.org/10.1007/s10742-024-00327-4.

[15] A. F. Hayes and M. Scharkow, "The relative trustworthiness of inferential tests of the indirect effect in statistical mediation analysis: does method really matter?," *Psychological science,* vol. 24, no. 10, pp. 1918-1927, 2013, https://doi.org/10.1177/0956797613480187.

[16] H. Zhang and X. Li, "A framework for mediation analysis with massive data," *Statistics and Computing,* vol. 33, no. 4, p. 86, 2023, https://doi.org/10.1007/s11222-023-10255-x.

[17] K. M. Marcoulides and T. Raykov, "Evaluation of Variance Inflation Factors in Regression Models Using Latent Variable Modeling Methods," (in eng), *Educ Psychol Meas,* vol. 79, no. 5, pp. 874-882, 2019, https://doi.org/10.1177/0013164418817803.

[18] J. Kidd *et al.*, "Hypothesis tests of indirect effects for multiple mediators: J. Kidd et al," *Statistical Methods & Applications,* vol. 34, no. 1, pp. 113-127, 2025, https://doi.org/10.1007/s10260-024-00777-7.

[19] S. Nemes *et al.*, "Integrative genomics with mediation analysis in a survival context," *Computational and Mathematical Methods in Medicine,* vol. 2013, no. 1, p. 413783, 2013, https://doi.org/10.1155/2013/413783.

[20] Y. Lin, Z. Guo, B. Sun, and Z. Lin, "Testing high-dimensional mediation effect with arbitrary exposure–mediator coefficients: Y. Lin et al," *Test,* pp. 1-50, 2025, https://doi.org/10.48550/arXiv.2310.05539.

https://doi.org/10.13057/ijas.v8i2.93376
E-ISSN: 2621-086X