Pak.j.stat.oper.res. Vol.17 No.1 2021 pp 273-289 DOI: http://dx.doi.org/10.18187/pjsor.v17i1.3148

Pakistan Journal of Statistics and Operation Research

Calculating non-centrality parameter for power analysis under structural equation modelling: An alternative

David Adedia[1*], Atinuke O. Adebanji[2], Simon Kojo Appiah[3]

*Corresponding Author

1. Department of Basic Sciences, University of Health and Allied Sciences, PMB 31, Ho, Ghana; Department of Statistics and Actuarial Science, Kwame Nkrumah University of Science and Technology, PMB, UPO, Kumasi, Ghana, dadedia@uhas.edu.gh
2. Department of Statistics and Actuarial Science, Kwame Nkrumah University of Science and Technology; KNUST-Laboratory for Interdisciplinary Statistical Analysis (LISA), Kwame Nkrumah University of Science and Technology, PMB, UPO, Kumasi, Ghana, aoadebanji.cos@knust.edu.gh
3. Department of Statistics and Actuarial Science, Kwame Nkrumah University of Science and Technology; KNUST-Laboratory for Interdisciplinary Statistical Analysis (LISA), Kwame Nkrumah University of Science and Technology, PMB, UPO, Kumasi, Ghana, skappiah@knust.edu.gh

Abstract

Identifying the most parsimonious model in structural equation modelling (SEM) is of utmost importance and the appropriate power estimation methods minimize the probabilities of Type I and Type II errors. The power of a test depends on the sample size, Type I error, degrees of freedom and effect size. The effect size choice for power analysis under SEM is so critical. The estimate of the population discrepancy function ($\hat{F_0}$), which is based on the sample discrepancy function, is usually used in power computations instead of the sample discrepancy function ($\hat{F}$). Although the sample discrepancy function measures the actual difference between the model implied covariance matrix and covariance matrix of the data, it was considered as a biased estimator to be used as an effect size for power analysis. In this study, a modified approach of using the sample discrepancy function as an effect size in calculating the non-centrality parameter for power is proposed. This is compared to the approach in MacCallum et al. (1996) at different degrees of freedom and sample size specifications — taken from 50 to 2000. The relative efficiency of $\hat{F}$ was derived, and its asymptotic unbiasedness for calculating the non-centrality parameter for power analysis was assessed. As the sample size and degrees of freedom increased, the difference between the power of a test for both methods reduced to zero. The results showed that the values for the power of a test are the same for the modified and traditional approaches for large sample sizes and degrees of freedom. The findings also revealed that the sample discrepancy function ($\hat{F}$) is asymptotically unbiased for power analysis.

Key Words: Structural equation modelling; Sample discrepancy function; Effect size; Root mean square error approximation; Non-centrality parameter; Power analysis.

Mathematical Subject Classification: 62G35, 62F35.

1. Introduction

The focus of structural equation modelling (SEM) is to find theoretical variables (constructs), and model relationships or discover a model that can explain the relationship between the variables under consideration (Yuan et al., 2015; Suhr, 2006) by explaining the covariance structure in the observed variables (Hox and Bechger, 1998). The SEM adopts a combination of factor, regression and path analyses. A confirmatory factor analysis, which is constrained by a number of factors and variables per factor (Cui, 2012; Suhr, 2008) is usually used with SEM. The structural model by Wright (1921), takes into consideration the path diagram, equations and type of effects (direct or indirect) one variable has on the other. The equations depict the relationship between covariances or correlations of the variables with the model parameters. The theoretical variables are computed in addition to the covariances existing between them, and the variances of the residuals of the variables. Hypothesized relationships are tested simultaneously with SEM (Hu and Bentler, 1999). The model is then assessed on how it fits the data. To minimize the probabilities of Type I and Type II errors in model fit assessment, power of a test is necessary. The power of a test depends on the sample size, Type I error, degrees of freedom and effect size. The effect size choice for power analysis under SEM is very critical. The estimate of the population discrepancy function ($\hat{F_0}$), which is based on the sample discrepancy function, is usually used in power computations instead of the sample discrepancy function ($\hat{F}$). Although the sample discrepancy function measures the actual difference between the model implied covariance matrix and covariance matrix of the data, it was considered as a biased estimator to be used as effect size for power analysis (MacCallum et al. 1996; Ryu 2014). In this study, a modified approach of using the sample discrepancy function as an effect size in calculating the non-centrality parameter for power is proposed. This is compared to the approach in MacCallum et al. (1996) at different degrees of freedom and sample size specifications. The distribution of $\hat{F_0}$ and the relative efficiency of $\hat{F}$ were derived. The study assessed the asymptotic unbiasedness of $\hat{F}$ in calculating the non-centrality parameter for power analysis.

2. Methodology

2.1. Structural Equation Model

The structural equation model and measurement models are represented respectively by the follows equations:

$$\eta = (1-\beta)^{-1}\Gamma\xi + (1-\beta)^{-1}\zeta \tag{1}$$

$$X = \Lambda_x \xi + \delta, \tag{2}$$

$$Y = \Lambda_y \eta + \epsilon, \tag{3}$$

where η contains the endogenous latent variables, ξ contains the exogenous latent variables, β contains the coefficients of η variables, ζ contains the random disturbances or errors associated with the structural model, Γ is the matrix of the coefficients of exogenous latent variables, ϵ and δ are the random errors associated with the measurement models for determining the endogenous and exogenous latent variables, respectively (Bollen, 1989). Therefore, the focus of the SEM is to estimate the parameters in the equations (1-3). To estimate the parameters, the function, $F(\theta)$, which minimizes $S - \Sigma(\theta)$, such that $F(\theta) \geq 0$ and small as possible (Bollen, 1989; Cui, 2012) is used, where θ is the parameter space for equations (1-3). Using the traditional maximum likelihood estimation based function, the parameters are estimated as follows. Let $\boldsymbol{Z}$ be a random variable for the structural model, such that $\boldsymbol{Z}_1, \boldsymbol{Z}_2, ..., \boldsymbol{Z}_N$ are independent. Given that $\boldsymbol{Z} \sim N_k(\mathbf{0}, \boldsymbol{\Sigma})$,

$$f(\boldsymbol{Z}_i) = \frac{1}{(2\pi)^{(k)/2}\,|\boldsymbol{\Sigma}|^{1/2}} \exp\left[-\frac{1}{2}\boldsymbol{Z}_i^T \boldsymbol{\Sigma}^{-1}\boldsymbol{Z}_i\right],\ i = 1, 2, ..., N.$$

Taking likelihood results in

$$L(\boldsymbol{\Sigma}(\boldsymbol{\theta})) = \prod_{i=1}^{N}(f(\boldsymbol{Z}_i)) = \frac{1}{(2\pi)^{N(k)/2}\,|\boldsymbol{\Sigma}(\boldsymbol{\theta})|^{N/2}} \exp\left[-\frac{1}{2}\sum_{i=1}^{N}\boldsymbol{Z}_i^T \boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\boldsymbol{Z}_i\right]$$

Taking natural log and considering the implied model $\boldsymbol{\Sigma}(\boldsymbol{\theta})$, we have

$$\begin{aligned}\ln L(\boldsymbol{\Sigma}(\boldsymbol{\theta})) &= -\frac{N(k)}{2}(2\pi) - \frac{N}{2}\ln|\boldsymbol{\Sigma}(\boldsymbol{\theta})| - \frac{1}{2}\sum_{i=1}^{N}(\boldsymbol{Z}_i^T\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\boldsymbol{Z}_i) \\ &= -\frac{N(k)}{2}(2\pi) - \frac{N}{2}\left[\ln|\boldsymbol{\Sigma}(\boldsymbol{\theta})| + \text{tr}\left(\boldsymbol{S}\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\right)\right]\end{aligned} \tag{4}$$

Supposing the $\boldsymbol{S}_{(k\times k)} \approx \boldsymbol{\Sigma}(\boldsymbol{\theta})_{(k\times k)}$, then equation (4) becomes

$$\begin{aligned}\ln L(\boldsymbol{S}) &= -\frac{N(k)}{2}(2\pi) - \frac{N}{2}\left[\ln|\boldsymbol{S}| + \text{tr}\left(\boldsymbol{S}\boldsymbol{S}^{-1}\right)\right] \\ &= -\frac{N(k)}{2}(2\pi) - \frac{N}{2}\left[\ln|\boldsymbol{S}| + (k)\right]\end{aligned} \tag{5}$$

In order to estimate the parameters, we seek a function $F(\boldsymbol{\theta})$ which minimizes $\boldsymbol{S} - \boldsymbol{\Sigma}(\boldsymbol{\theta})$, such that $F(\boldsymbol{\theta}) \approx \min(\boldsymbol{S} - \boldsymbol{\Sigma}(\boldsymbol{\theta}))$ (1, 3).

$$\begin{aligned}F(\boldsymbol{\theta}) &= \ln L(\boldsymbol{S}) - \ln L[\boldsymbol{\Sigma}(\boldsymbol{\theta})] \\ &= -\frac{N}{2}\left[\ln|\boldsymbol{S}| + (k)\right] + \frac{N}{2}\left[\ln|\boldsymbol{\Sigma}(\boldsymbol{\theta})| + \text{tr}\left(\boldsymbol{S}\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\right)\right] \\ &= \frac{N}{2}\left[\ln|\boldsymbol{\Sigma}(\boldsymbol{\theta})| + \text{tr}\left(\boldsymbol{S}\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\right) - \ln|\boldsymbol{S}| - k\right],\end{aligned} \tag{6}$$

where k is the number of variables in the structural equation model, Σ is the covariance matrix for the population which we estimate using sample covariance matrix S, and $\Sigma(\theta)$ is the model implied covariance matrix. Equation (6) is non-linear so iterative approach such as Newton-Raphson's is employed in the minimization (Deng et al., 2018).

2.2. Model Adequacy Test

The likelihood ratio (omnibus) test is usually undertaken in SEM for testing whether $\Sigma = \Sigma(\theta)$ or not (Kelloway, 1995). This test is distributed as χ^2_{df}, with degrees of freedom $(df) = \dfrac{k(k+1)}{2} - t$, where t is the number of parameters to be estimated (MacCallum et al., 1996). A non-significance of this test implies a non-significant discrepancy between these two covariance matrices. This shows that the proposed relationships in the model can provide the population covariance matrix (Kelloway, 1995; Bollen, 1989; Cui, 2012). The chi-square test assesses the hypothesis that all the residuals are zero with test statistic, $\chi^2 = (N-1)F(\theta)$. This exact fit test with large sample size rejects the null hypothesis and the test statistic from small sample size lacks power. As a result, other methods of model assessment were proposed including close fit and not close fit test, which perform the similar omnibus test as chi-square by using root mean square error approximation (Hooper et al., 2008; MacCallum et al., 1996).

2.3. Root Mean Square Error Approximation

The root mean square error approximation (RMSEA) is a model-data fit index (Ryu, 2014; Cui, 2012; Iacobucci, 2010), which is given by:

$$RMSEA = \sqrt{\frac{F_0}{df}} = \sqrt{max\left[\left(\frac{\chi^2 - df}{df(N-1)}\right), 0\right]}. \tag{7}$$

where F_0 is the population discrepancy function. A $RMSEA$ of zero (0) implies a good fit, whilst a value greater than 0 indicates a bad fit. Studies show that a model is considered close fit if $RMSEA < 0.05$, an average fit if $0.05 \le RMSEA \le 0.08$, mediocre fit if $0.08 < RMSEA \le 0.10$ and a poor fit if $RMSEA > 0.10$ (Cui, 2012; Hooper et al., 2008).

2.3.1. Power of a Test as a Fit Index

The power of a test is the probability of correctly rejecting a false null hypothesis. In structural equation modelling and confirmatory factor analysis, rejecting a null hypothesis of a model fit implies the hypothesized model does not fit the data. Initially, testing the hypothesis $H_0 : \Sigma = \Sigma(\theta)$ as against $H_1 : \Sigma \neq \Sigma(\theta)$, which is the exact fit test (or $H_0 : \epsilon = 0$ as against $H_1 : \epsilon \neq 0$) was found to be rejecting good models with large sample sizes (MacCallum et al., 1996; Browne and Cudeck, 1993). To avoid this, Browne and Cudeck (1993) developed the hypothesis testing of close fit ($H_0 : \epsilon \leq 0.05$) whilst MacCallum et al. (1996) worked on not close fit ($H_0 : \epsilon \geq 0.05$). They computed the power of a test by determining the non-centrality parameter for the null and alternative hypotheses using the RMSEA by Steiger and Lind (1980). The power of a test for close fit, when $\epsilon_0 < \epsilon_1$ is given by $P(\chi^2_{(df,\lambda_1)} \geq \chi^2_{critical})$, where $\chi^2_{(df,\lambda_1)}$ is the distribution of the test statistic with degrees of freedom, df, non-centrality parameter λ_1, and ϵ_0 and ϵ_1 are the hypothesized RMSEA values for the null and alternative hypotheses respectively. Given that the H_0 is true, $\chi^2_{critical}$ becomes the reliability coefficient (critical value) for the test statistic. However, if $\epsilon_0 > \epsilon_1$, we test a hypothesis of not close fit and the power of a test is determined by $P(\chi^2_{(df,\lambda_1)} \leq \chi^2_{critical})$.

In testing hypothesis of exact fit, $H_0 : \Sigma = \Sigma(\theta)$, the test statistic $(N-1)\hat{F}$ is used. This approach of assessing a model fit was considered by Browne and Cudeck (1993) and MacCallum et al. (1996) as being too strict, so the suggestion to the testing of close fit was provided and so also was not close fit. Non-central chi-square distribution has been used in this procedure by the researchers in this area of study (Cui, 2012). A test statistic is said to be distributed as non-central chi-square if the expectation is given by the addition of its degrees of freedom and a non-centrality parameter λ. Given $H_0 : \Sigma = \Sigma(\theta)$, if $\hat{F} = 0$, then the test statistic $\chi^2 = (N-1)\hat{F}$ follows central chi-square distribution. However, in real-life applications $\hat{F} \neq 0$. As a result, $\chi^2 = (N-1)\hat{F}$ will follow a non-central χ^2 with degrees of freedom df and non-centrality parameter $\lambda = (N-1)F_0$. According to MacCallum et al. (1996), given the population RMSEA $(\epsilon) = \sqrt{\frac{F_0}{df}}$, we could estimate it using $\hat{\epsilon} = \sqrt{\frac{\hat{F}_0}{df}}$, where $\hat{F}_0 = \hat{F} - \frac{df}{N-1}$. Now $\hat{F}_0$ could be expressed as a function of $\hat{\epsilon}$, given by $\hat{F}_0 = \hat{\epsilon}^2(df)$, which implies that

$$\lambda = (N-1)\hat{F}_0 = (N-1)\hat{\epsilon}^2(df) \tag{8}$$

where ϵ is function of F_0 which is unknown.

$\hat{F}$, which measures the actual difference between the sample and implied covariance matrices is not used as an effect size for determining the non-centrality parameter because it was considered biased (MacCallum et al. 1996; Ryu 2014). To assess the asymptotic unbiasedness of $\hat{F}$ to be used in determining the non-centrality parameter for power analysis, the non-centrality parameter is determined based on $\hat{F}$ as follows. In practice, we estimate the ϵ using RMSEA as in equation (9).

$$\hat{\epsilon} = \sqrt{\frac{\chi^2 - df}{df(N-1)}} = \sqrt{\frac{\hat{F}(N-1) - df}{df(N-1)}} = \sqrt{\frac{\hat{F}}{df} - \frac{1}{N-1}} \tag{9}$$

The covariance structure modelling usually considers large sample sizes to satisfy asymptotic assumptions. So as $N \Longrightarrow \infty$,

$$\hat{\epsilon} = \sqrt{\frac{\hat{F}_0}{df}} = \sqrt{\frac{\hat{F}}{df} - \frac{1}{N-1}} \approx \sqrt{\frac{\hat{F}}{df}} \tag{10}$$

For a large sample data, we could determine λ using $\hat{F}$ rather than $\hat{F}_0$, where $\lambda = (N-1)\hat{F}$, and from equation (9), $\hat{F} = \left(d\hat{f}\epsilon^2 + \frac{df}{N-1}\right)$, where $\frac{df}{N-1}$ cannot easily go to zero since the degrees of freedom depends on the model specification. This implies that the non-centrality parameter $\lambda = (N-1)F_0$ will now become

$$\lambda = (N-1)\left(df\epsilon^2 + \frac{df}{N-1}\right) = (N-1)\hat{\epsilon}^2 df + df. \tag{11}$$

Hence, given sample data, we can estimate $\hat{F}$. MacCallum and Hong (1997) also used the goodness of fit index (GFI) and adjusted goodness of fit index (AGFI) as effect size for deriving the power of a test. Kim (2005) computed the power of a test for the fit index by McDonald (1989), comparative fit index (CFI), Steiger's gamma and RMSEA by determining the non-centrality parameters based on them. However, apart from RMSEA, other fit indices showed some limitations to be used for power analysis (MacCallum and Hong, 1997; Cui, 2012; Kim, 2005). For the RMSEA-based approach, Cui (2012) derived the mean and variance of the sample discrepancy function and determined its distribution.

2.3.2. Distribution of the Discrepancy Functions $(\hat{F})$ and $(\hat{F}_0)$

The distribution of the population estimate of the discrepancy is derived to provide the relative efficiency of the sample discrepancy function. The two discrepancy functions were presented as choices for the effect size in deriving the non-centrality parameter for power analysis in SEM. For a true null hypothesis,

$$\chi^2 = (N-1)\hat{F}, \tag{12}$$

which follows a central chi-square (χ_c^2) with mean df_c and variance $2df_c$: $\chi^2 \sim \chi_c^2(df_c, 2df_c)$. However, a false null hypothesis will cause $\chi^2 = (N-1)\hat{F}$ to follow a non-central chi-square (χ_{nc}^2) with mean $df_{nc} + \lambda$ and variance $2df_{nc} + \lambda$: $\chi^2 \sim \chi_{nc}^2(df_{nc} + \lambda, 2df_{nc} + \lambda)$ (Cui, 2012). For covariance structure models, $\hat{F}$ depicts the level of misspecification in the model and could be used as the effect size instead of $\hat{\lambda} = \chi_{nc}^2 - df_{nc}$ for sample data. From equation (12),

$$\hat{F} = \frac{\chi^2}{N-1}, \tag{13}$$

of which for a true null hypothesis, the mean, $E(\hat{F}) = E\left(\frac{\chi^2}{N-1}\right) = \frac{df_c}{N-1}$, and variance, $Var(\hat{F}) = Var\left(\frac{\chi^2}{N-1}\right) = \frac{2df_c}{(N-1)^2}$. For a false null hypothesis, the mean $E(\hat{F}) = E\left(\frac{\chi^2}{N-1}\right) = \frac{df_{nc}+\lambda}{N-1}$, and variance, $Var(\hat{F}) = Var\left(\frac{\chi^2}{N-1}\right) = \frac{2df_{nc}+4\lambda}{(N-1)^2}$. Rewriting equation (13) as $\hat{F} = \frac{1}{N-1}\chi^2$, $\frac{1}{N-1}$ is a constant hence, $\hat{F}$ has similar features as the χ^2 distribution (Cui, 2012). Therefore, from above, for a true null hypothesis, $\hat{F}$ is chi-square distributed with mean and variance as $\hat{F} \sim \chi^2\left(\frac{df_c}{N-1}, \frac{2df_c}{(N-1)^2}\right)$. Also, for a false null hypothesis, $\hat{F}$ is chi-square distributed with mean and variance as $\hat{F} \sim \chi^2\left(\frac{df_{nc}+\lambda}{N-1}, \frac{2df_{nc}+4\lambda}{(N-1)^2}\right)$.

For the less biased estimator of the population discrepancy function $(\hat{F}_0)$, the mean and variance are derived under a true null hypothesis as

$$\begin{aligned} E(\hat{F}_0) &= E\left(\hat{F} - \frac{df_c}{N-1}\right) \\ &= \frac{df_c}{N-1} - \frac{df_c}{N-1} = 0 \end{aligned} \tag{14}$$

$$\begin{aligned} Var(\hat{F}_0) &= Var\left(\hat{F} - \frac{df_c}{N-1}\right) \\ &= \frac{2df_c}{(N-1)^2} - 0 \\ &= \frac{2df_c}{(N-1)^2}, \end{aligned} \tag{15}$$

and for the false null hypothesis, the mean and variance are derived as

$$\begin{aligned} E(\hat{F}_0) &= E\left(\hat{F} - \frac{df_{nc}}{N-1}\right) \\ &= \frac{df_{nc}+\lambda}{N-1} - \frac{df_{nc}}{N-1} \\ &= \frac{\lambda}{N-1} \end{aligned} \tag{16}$$

$$\begin{aligned} Var(\hat{F}_0) &= Var\left(\hat{F} - \frac{df_{nc}}{N-1}\right) \\ &= \frac{2df_{nc}+4\lambda}{(N-1)^2} - 0 \\ &= \frac{2df_{nc}+4\lambda}{(N-1)^2}. \end{aligned} \tag{17}$$

$\hat{F}_0$ is chi-squared distributed with mean 0 and variance $\frac{2df_c}{(N-1)^2}$ under a true null hypothesis and chi-squared distributed with mean $\lambda/(N-1)$ and variance $\frac{2df_{nc}+4\lambda}{(N-1)^2}$ for a false null hypothesis. $\hat{F}_0$ and $\hat{F}$ have different means but the same variance when the null hypothesis is true, and when it is false. Thus, the relative efficiency of the sample discrepancy function with respect to the less biased estimator is 1:

$$\frac{Var(\hat{F}_0)}{Var(\hat{F})} = \frac{\frac{2df_c}{(N-1)^2}}{\frac{2df_c}{(N-1)^2}} = 1 \tag{18}$$

for a true null hypothesis and a false null hypothesis, the relative efficiency is also 1:

$$\frac{Var(\hat{F}_0)}{Var(\hat{F})} = \frac{\frac{2df_{nc}+4\lambda}{(N-1)^2}}{\frac{2df_{nc}+4\lambda}{(N-1)^2}} = 1. \tag{19}$$

These show that both estimators of the population discrepancy function are equally efficient.

2.3.3. Simulation Study

The methodology was implemented on the parameters from two simulated datasets. The first dataset was simulated on 11 manifest variables (M) with size $N = 50$ to measure 3 theoretical variables (T) leading to the model as illustrated in Figure 1. The same model (Figure 1) was used to simulate another dataset of size $N = 100$. The parameters from the modelled first dataset were used for the power of a test of not close fit and the parameters from the modelled second dataset were used for the power of a test of close fit test. For the power analysis, the RMSEA and degrees of freedom from each dataset were used with sample sizes from 50 to 2,000 to ascertain the behaviour of the modified approach with the traditional one. We also varied the alternative hypothesized values for the close fit test (0.09, 0.08 and 0.07) and for not close fit test (0.01, 0.02 and 0.03) with the same degrees of freedom of 38 and determined power of a test based on them.

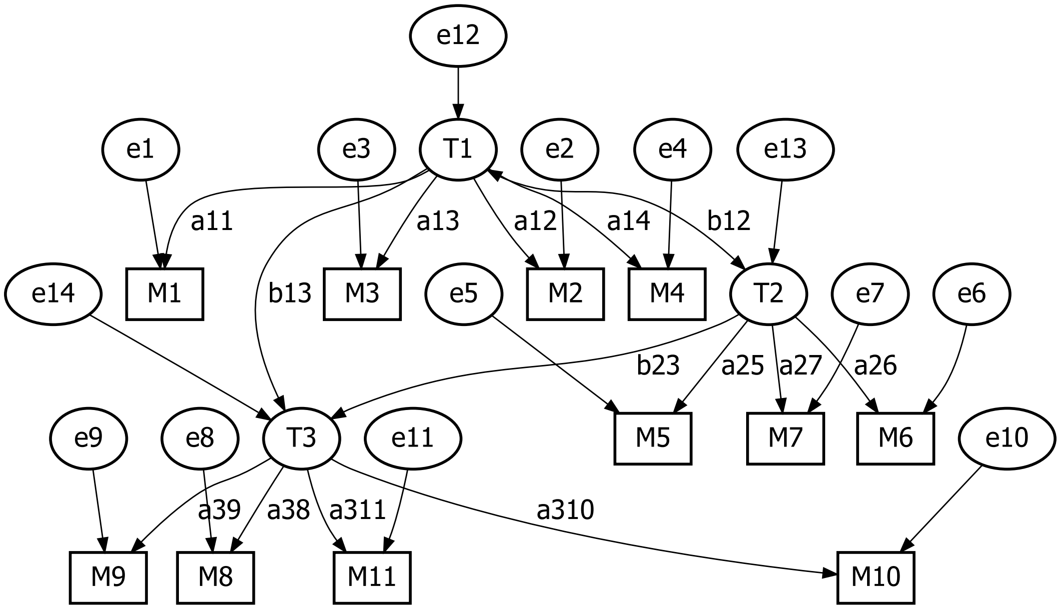


Figure 1: Path diagram for model used for data generation

3. Results and Discussion

The derived RMSEA and degrees of freedom for the two data sets are in Table 1.

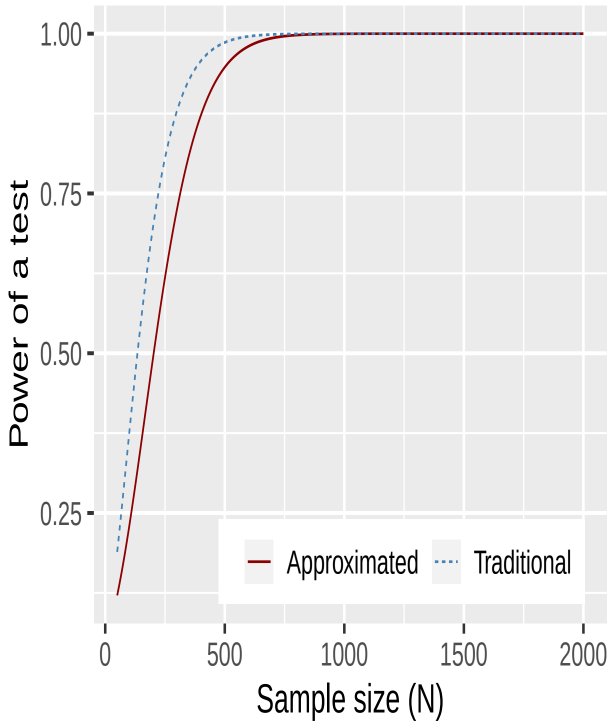
Table 1: RMSEA and degrees of freedom for the two simulated datasets

Fit measures	Data set 1 (N=50)	Data set 2 (N=100)
RMSEA	0.019	0.08
Degrees of freedom	41	41

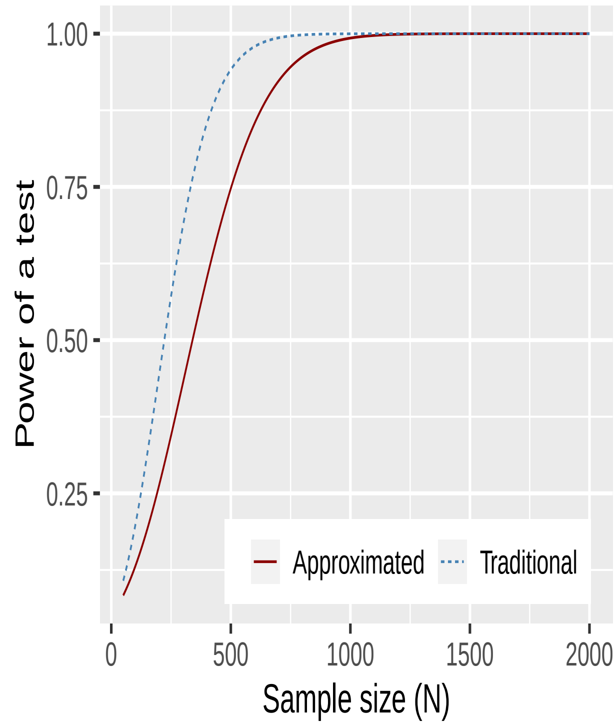
Figure 2 and Figure 3 are plotted using the parameters from the two simulated datasets. Figure 2 contains the power plots for both traditional and approximated approaches. The power plots are the close fit in subfigure (a) ($\hat{\epsilon} = 0.08$, $df = 41$ and $N = 50 : 2000$); not close fit in subfigure (b) ($\hat{\epsilon} = 0.019$, $df = 41$ and $N = 50 : 2000$); close fit in subfigure (c) ($\hat{\epsilon} = 0.08$, $df = 1 : 400$ and $N = 500$) and not close fit in subfigure (d) ($\hat{\epsilon} = 0.019$, $df = 1 : 400$ and $N = 500$). The differences in the power of a test for both approaches are respectively plotted in Figure 3: subfigure (a) to (d). Figures 4 - 9 are plotted using alternative hypothesized values of 0.09, 0.08 and 0.07 for close fit test; 0.01, 0.02 and 0.03 for not close fit test and degrees of freedom of 38 for both tests. The subfigures are arranged as specified above but with stated alternative hypothesized values and degrees of freedom.

From subfigure (a) of Figure 2, the power of a test approached one (1) when the sample sizes were greater than 500. The power analysis in this close fit test showed that the power increases at a faster rate for the traditional method. It is clear from subfigure (a) of Figure 2 that power of a test increases with sample size as discussed by other researchers including MacCallum et al. (1996) and Cui (2012). This figure also showed that for smaller sample sizes, the modified method has smaller power of a test but as the sample size increases both approaches showed similar values for the power of a test. Subfigure (a) in Figure 3 showed an increasing difference in the values of the power of a test for the traditional and modified for sample sizes less than 250, but begins to reduce afterwards for sample sizes larger than that. Subfigure (b) of Figure 2 also showed that the power of a test for not close fit test increases with an increase in sample size even though not at a faster rate as that of close fit.

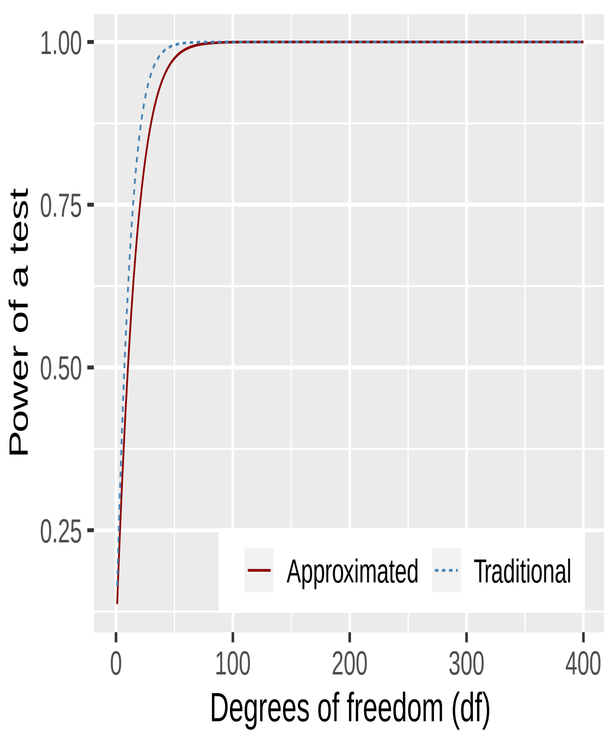
 DOI: http://dx.doi.org/10.18187/pjsor.v17i1.3148



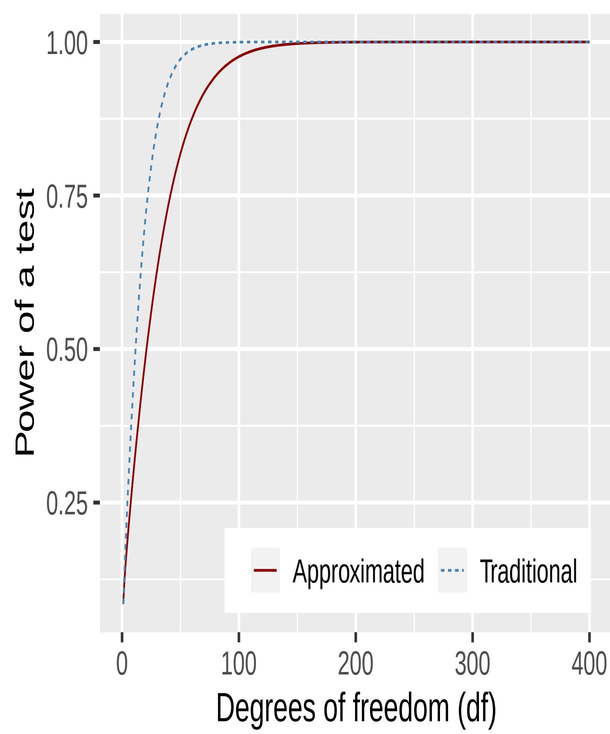
(a) Close fit test: $\epsilon_1 = 0.08$



(b) Not close fit test: $\epsilon_1 = 0.019$

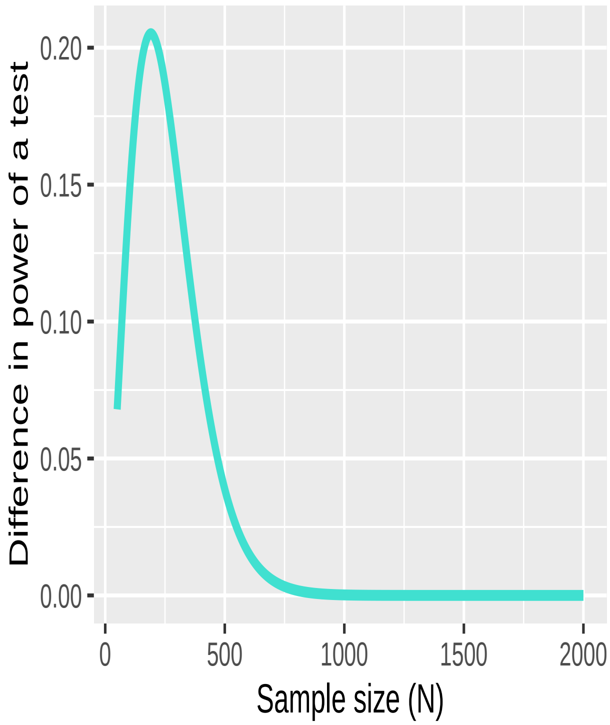


(c) Close fit test: $\epsilon_1 = 0.08$

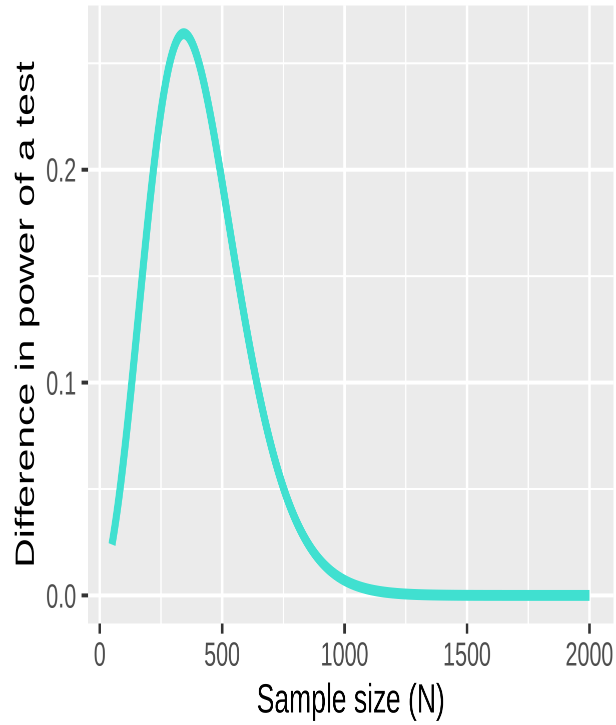


(d) Not close fit test: $\epsilon_1 = 0.019$

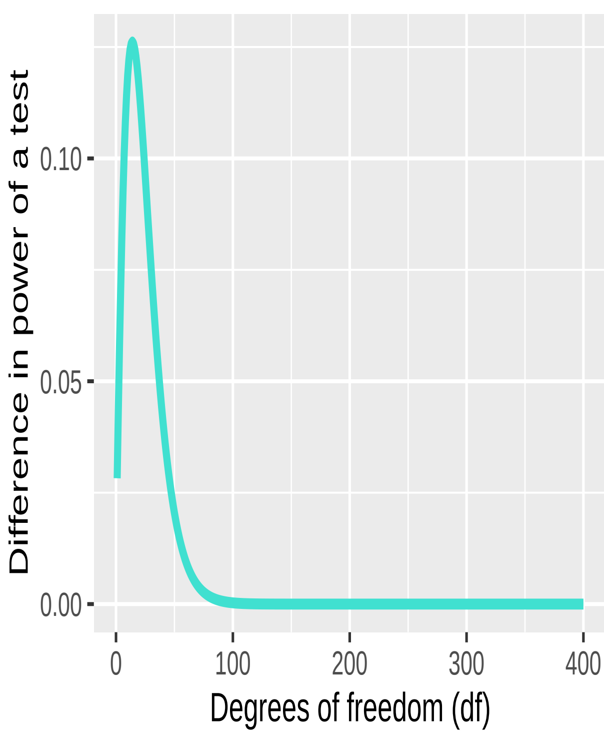
Figure 2: The power of a test for different sample sizes and degrees of freedom



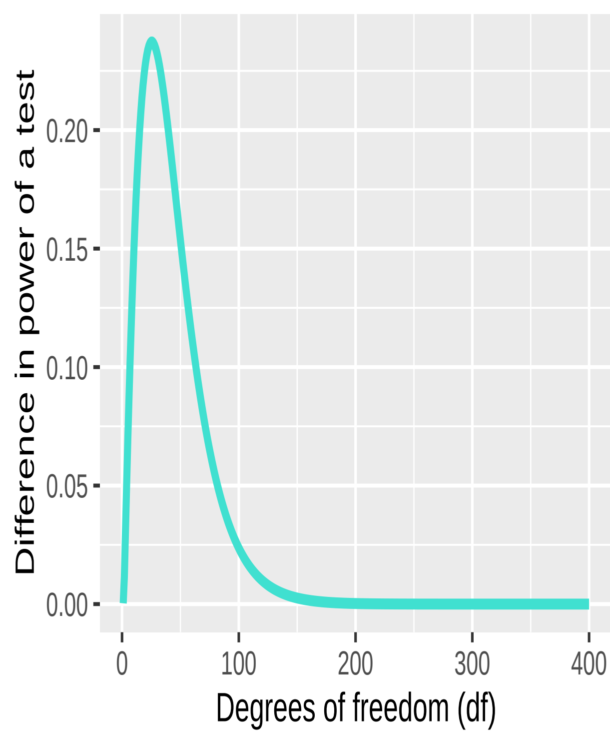
(a) Close fit test: $\epsilon_1 = 0.08$



(b) Not close fit test: $\epsilon_1 = 0.019$



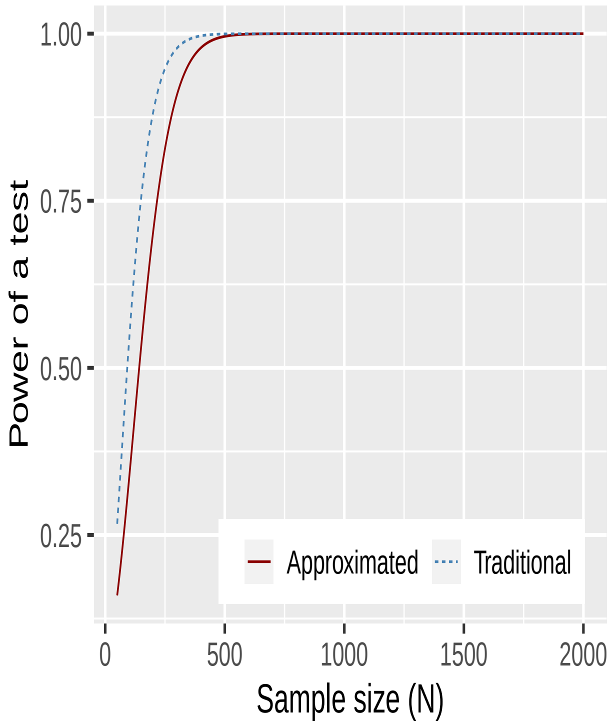
(c) Close fit test: $\epsilon_1 = 0.08$



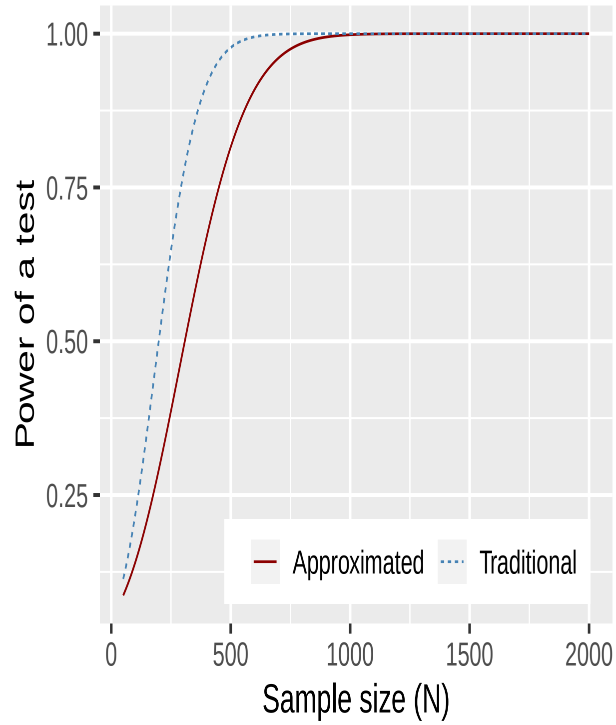
(d) Not close fit test: $\epsilon_1 = 0.019$

Figure 3: The difference in power of a test

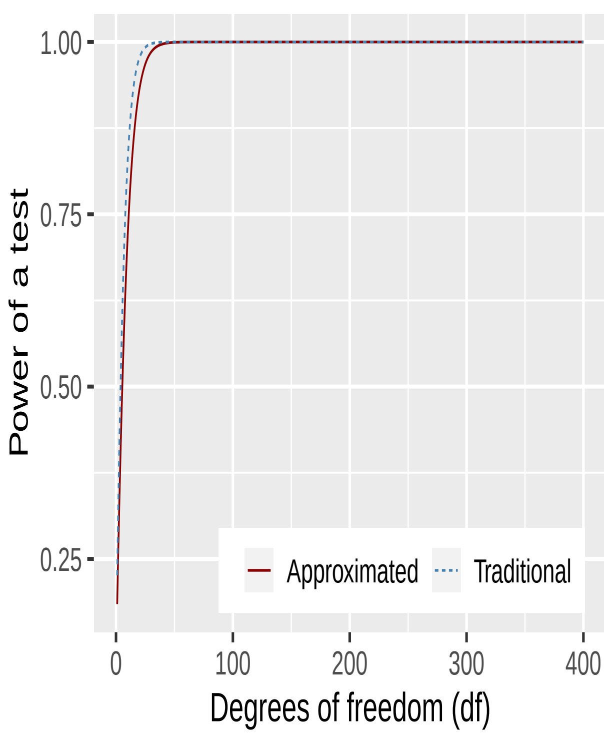
 DOI: http://dx.doi.org/10.18187/pjsor.v17i1.3148



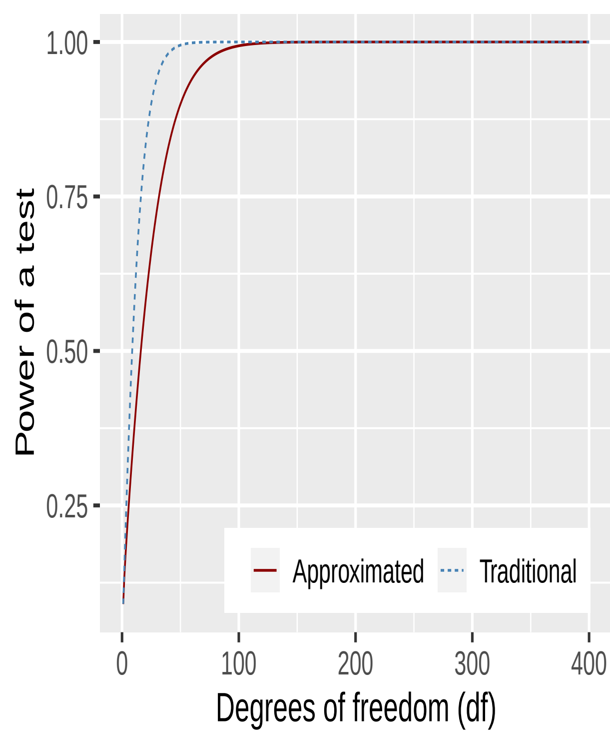
(a) Close fit test: $\epsilon_1 = 0.09$



(b) Not close fit test: $\epsilon_1 = 0.01$



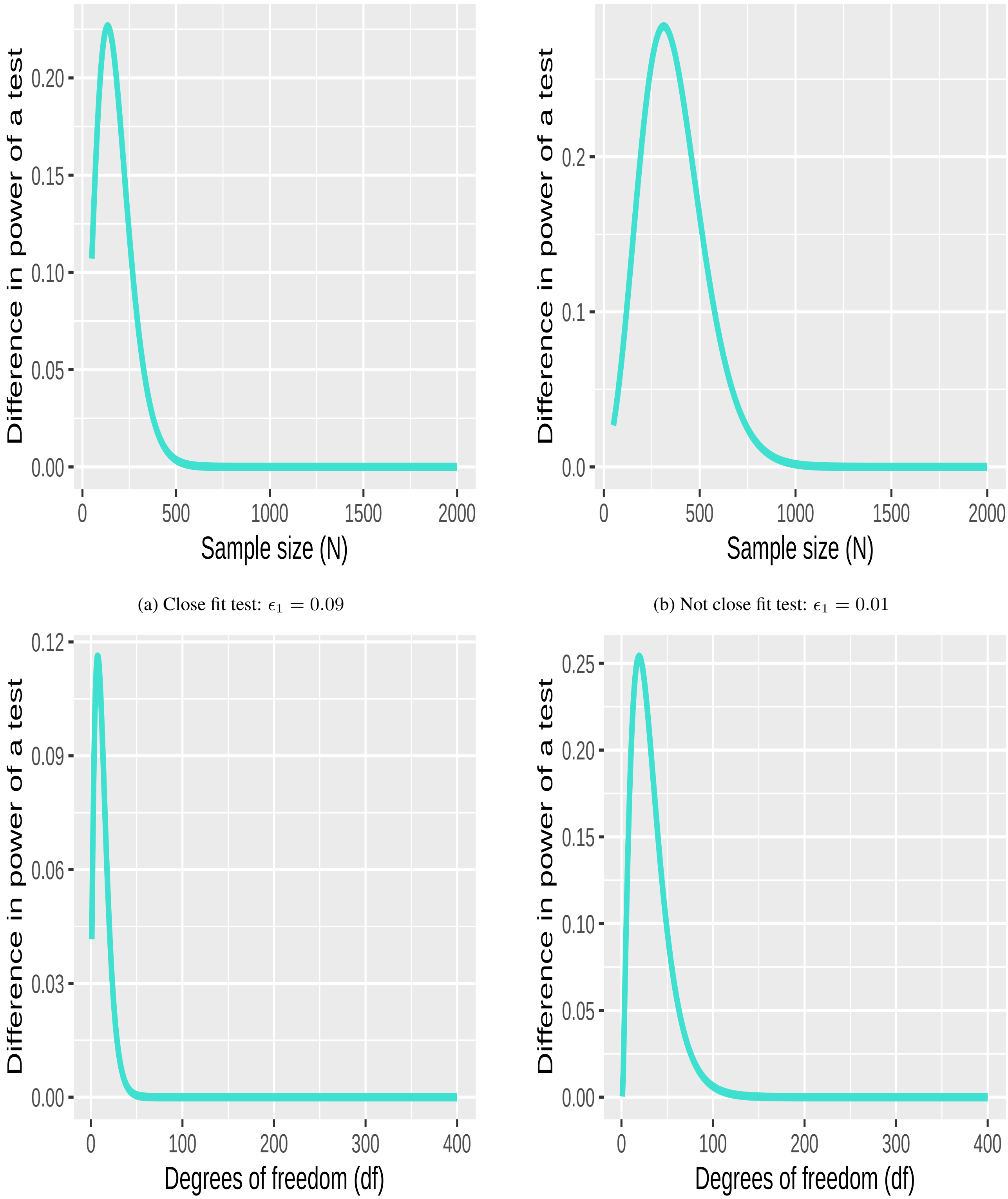
(c) Close fit test: $\epsilon_1 = 0.09$



(d) Not close fit test: $\epsilon_1 = 0.01$

Figure 4: The power of a test for different sample sizes and degrees of freedom

DOI: http://dx.doi.org/10.18187/pjsor.v17i1.3148



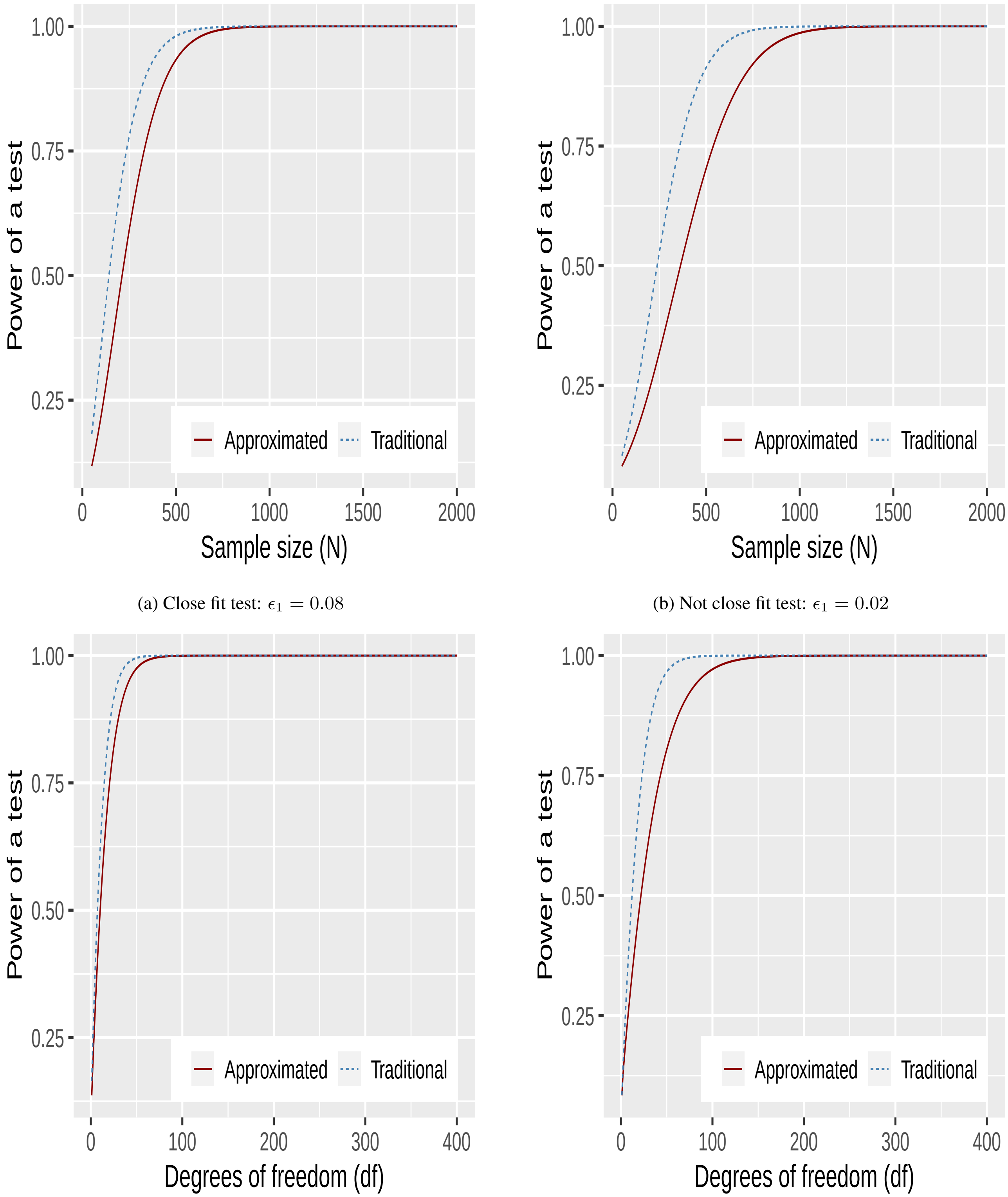
(a) Close fit test: $\epsilon_1 = 0.09$

(b) Not close fit test: $\epsilon_1 = 0.01$

(c) Close fit test: $\epsilon_1 = 0.09$

(d) Not close fit test: $\epsilon_1 = 0.01$

Figure 5: The difference in power of a test



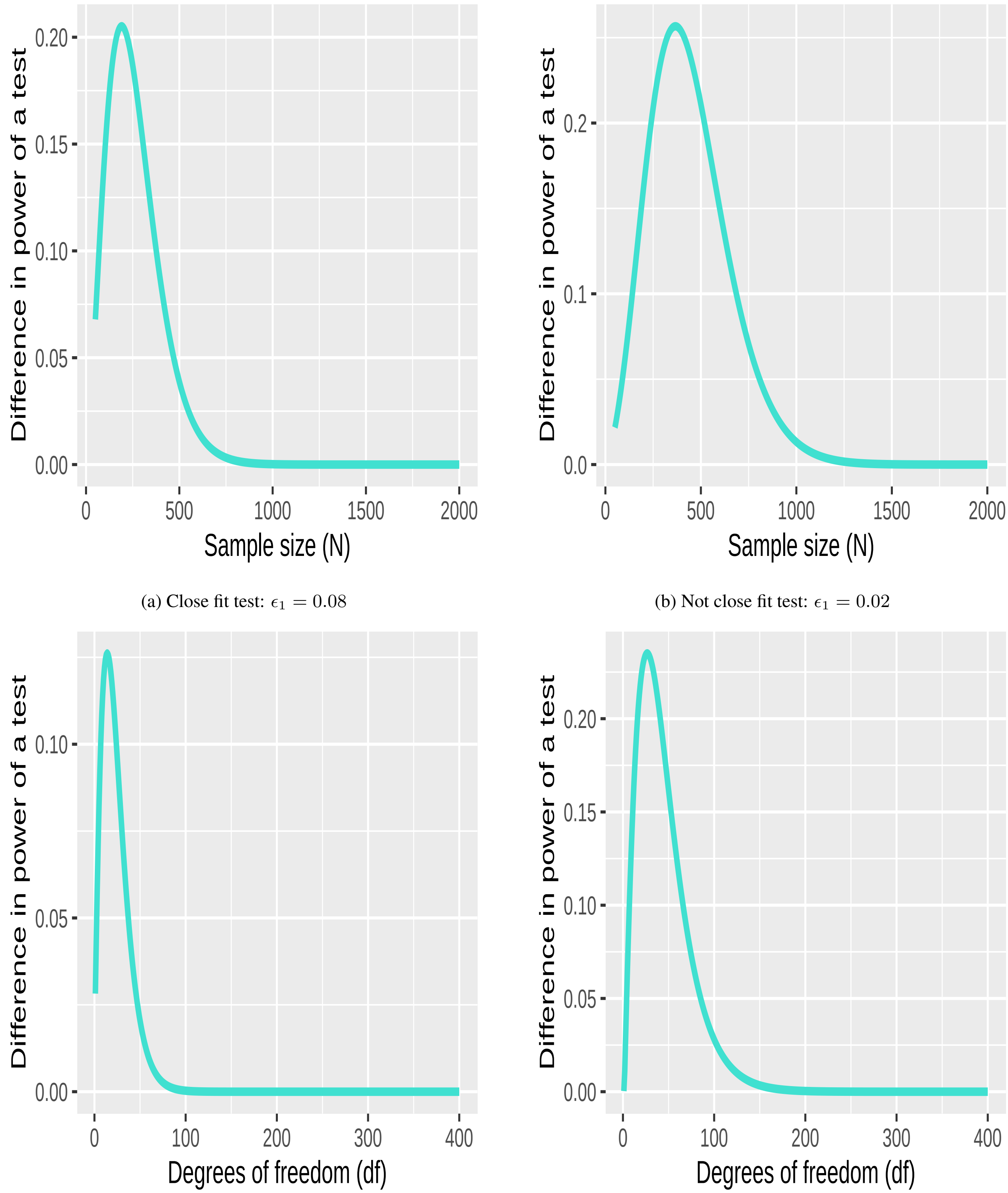
(a) Close fit test: $\epsilon_1 = 0.08$

(b) Not close fit test: $\epsilon_1 = 0.02$

(c) Close fit test: $\epsilon_1 = 0.08$

(d) Not close fit test: $\epsilon_1 = 0.02$

Figure 6: The power of a test for different sample sizes and degrees of freedom



(a) Close fit test: $\epsilon_1 = 0.08$

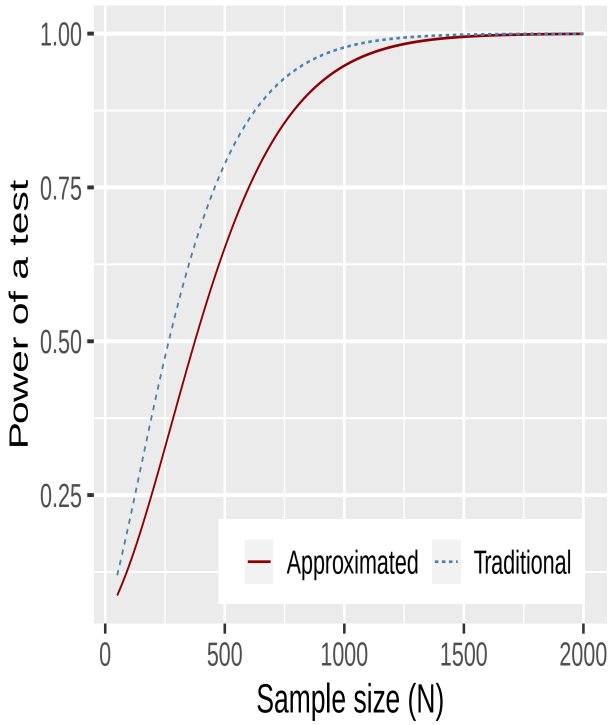
(b) Not close fit test: $\epsilon_1 = 0.02$

(c) Close fit test: $\epsilon_1 = 0.08$

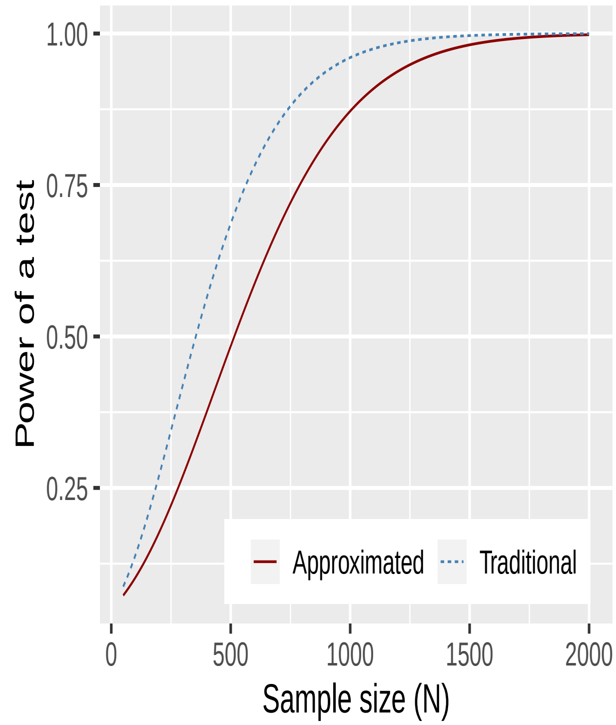
(d) Not close fit test: $\epsilon_1 = 0.02$

Figure 7: The difference in power of a test

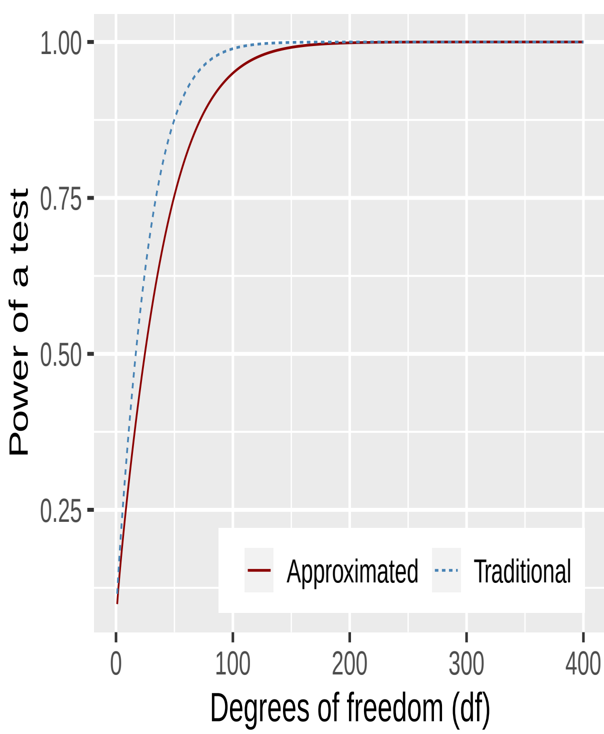
 DOI: http://dx.doi.org/10.18187/pjsor.v17i1.3148



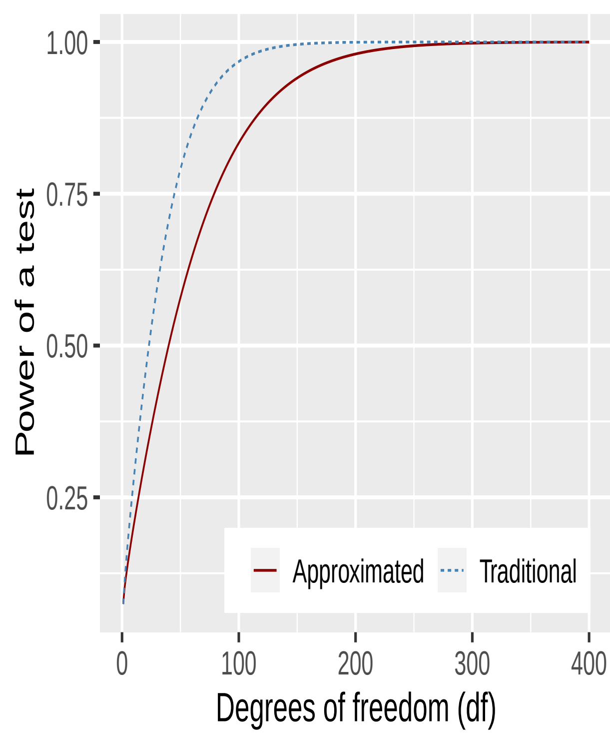
(a) Close fit test: $\epsilon_1 = 0.07$



(b) Not close fit test: $\epsilon_1 = 0.03$



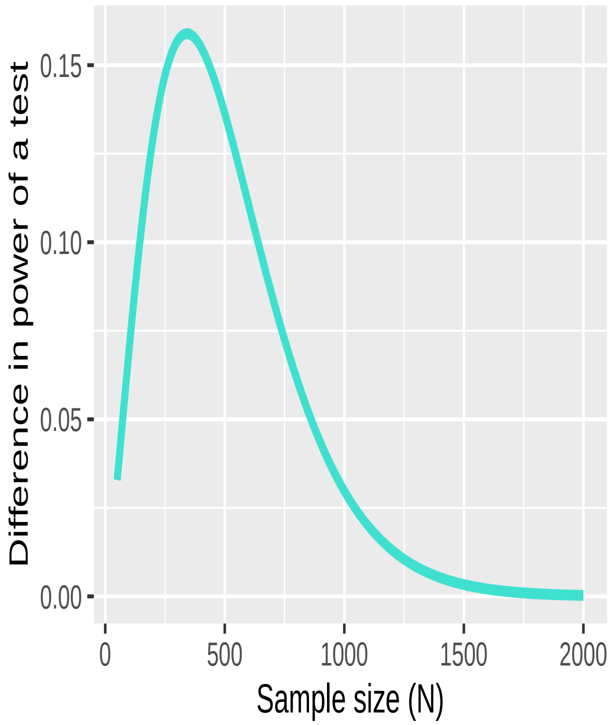
(c) Close fit test: $\epsilon_1 = 0.07$



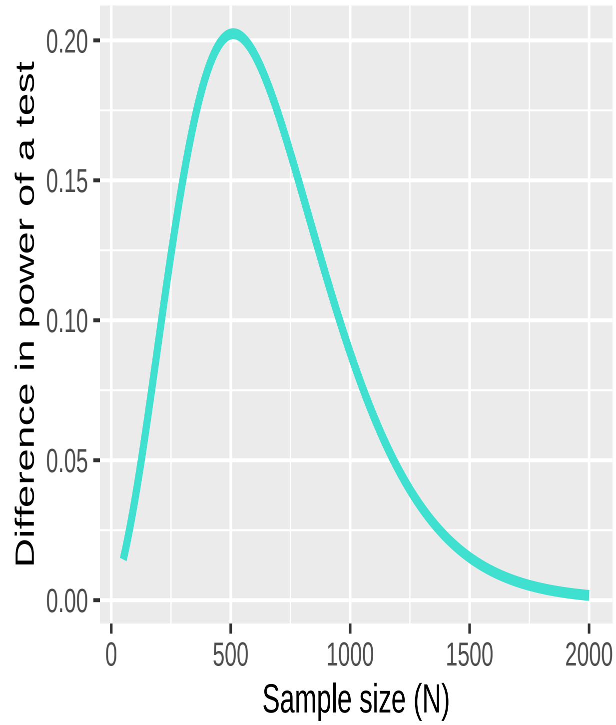
(d) Not close fit test: $\epsilon_1 = 0.03$

Figure 8: The power of a test for different sample sizes and degrees of freedom

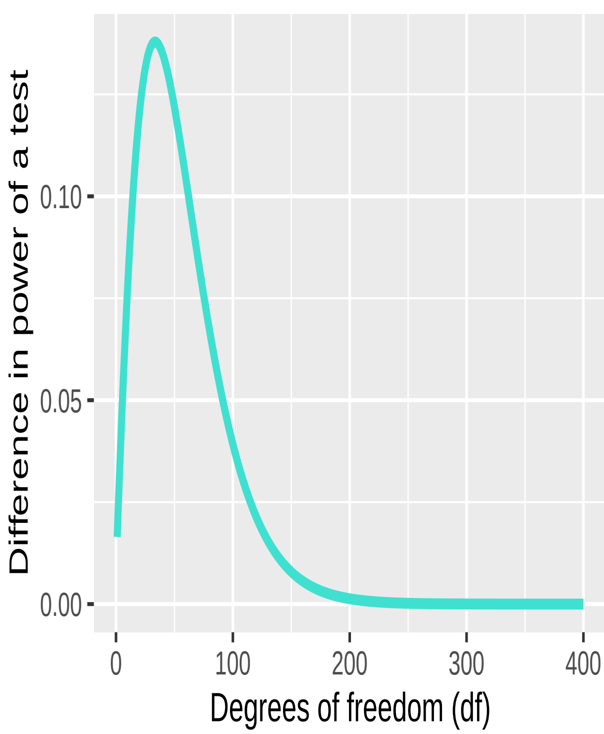
 DOI: http://dx.doi.org/10.18187/pjsor.v17i1.3148



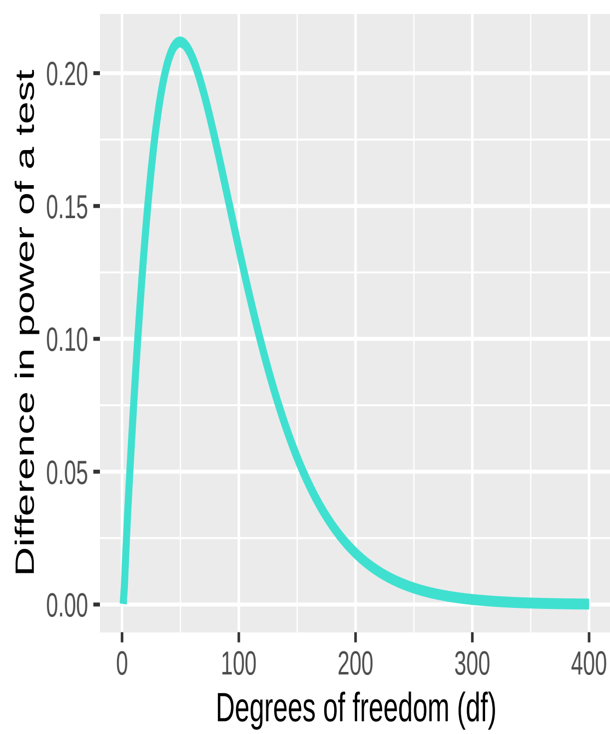
(a) Close fit test: $\epsilon_1 = 0.07$



(b) Not close fit test: $\epsilon_1 = 0.03$



(c) Close fit test: $\epsilon_1 = 0.07$



(d) Not close fit test: $\epsilon_1 = 0.03$

Figure 9: The difference in power of a test

The study also confirmed that higher effect size leads to a higher power and as a result higher root mean square error approximation for close fit also reports higher power of a test. Subfigure (b) of Figure 3 also showed an increase followed by a decrease in the difference in the values of the power of a test for the traditional and approximated but at a slower rate than subfigure (a) of Figure 3. Subfigure (c) and (d) of Figure 2 revealed a positive relationship between degrees of freedom and power of a test. The powers for both approaches were the same for large degrees of freedom, the differences in powers (subfigure (c) and (d) of Figure 3) also exhibited similar shapes for different sample sizes. The relationship that power of a test had with sample size and degrees of freedom as in Figure 2 is the same for Figures 4, 6 and 8; and also the difference in power of a test for both approaches in Figure 3 were similar in Figures 5, 7 and 9.

4. Conclusion

For close fit tests, the power of a test approached one (1) as the sample size and degrees of freedom increases but at a faster rate than the not close fit test. The traditional approach results in powers which approached one (1) faster than the approximated due to differences in non-centrality parameters. Moreover, the two approaches result in similar power of a test as the sample size increases. The power of a test also increases for a not close fit with increasing sample size and degrees of freedom. The power of a test increases with increase in distance between the hypothesized RMSEA values (ϵ_0 and ϵ_1), for null and alternative hypotheses for both approaches respectively. As the distance between ϵ_0 and ϵ_1 reduces, large sample size (N) and degrees of freedom (df) are required for high powers. Furthermore, the results show that $\hat{F}$ approaches $\hat{F}_0$ as the sample size increases. The traditional approach is better for all sample sizes and we recommend the use of the approximated approach for large sample datasets. We conclude that although $\hat{F}_0$ was referred to as a less biased estimator than $\hat{F}$ by MacCallum et al. (1996) and Ryu (2014), asymptotically, they are all unbiased. Moreover, these estimators have the same variance implying the same efficiency. Therefore, researchers in SEM can use $\hat{F}$ to determine the non-centrality parameter for power analysis with large samples.

References

1. Bollen, K. A. (1989). *Structural Equations with Latent Variables*. John Wiley & Sons, Inc.
2. Browne, M. W. and Cudeck, R. (1993). Alternative ways of assessing model fit. *Sage focus editions*, 154:136–136.
3. Cui, M. (2012). Effect-Size Index for Evaluation of Model- Data Fit in Structural Equation Modeling. Master's thesis, Florida State University.
4. Deng, L., Yang, M., and Marcoulides, K. M. (2018). Structural Equation Modeling With Many Variables: A Systematic Review of Issues and Developments. *Frontiers in psychology*, 9.
5. Hooper, D., Coughlan, J., and Mullen, M. (2008). Structural Equation Modelling : Guidelines for Determining Model Fit Structural equation modelling : guidelines for determining model fit. *Electronic Journal of Business Research Methods*, 6(1):53–60.
6. Hox, J. J. and Bechger, T. (1998). An Introduction to Structural Equation Modeling. *Family Science Review*, 11:354–373.
7. Hu, L.-t. and Bentler, P. M. (1999). Cutoff Criteria for Fit Indexes in Covariance Structure Analysis : Conventional Criteria Versus New Alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1):1–55.
8. Iacobucci, D. (2010). Structural equations modeling: Fit Indices, sample size, and advanced topics. *Journal of Consumer Psychology*, 20(1):90–98.
9. Kelloway, E. K. (1995). Structural equation modelling in perspective. *Journal of Organizational behavior*, 16(December 1994):215–224.
10. Kim, K. H. (2005). The relation among fit indexes, power, and sample size in structural equation modeling. *Structural Equation Modeling*, 12(3):368–390.
11. MacCallum, R. C., Browne, M. W., and Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological methods*, 1(2):130.
12. MacCallum, R. C. and Hong, S. (1997). Power analysis in covariance structure modeling using gfi and agfi. *Multivariate Behavioral Research*, 32(2):193–210.
13. McDonald, R. P. (1989). An index of goodness-of-fit based on noncentrality. *Journal of classification*, 6(1):97–103.
14. Ryu, E. (2014). Model fit evaluation in multilevel structural equation models. *Frontiers in Psychology*, 5:1–9.

15. Steiger, J. H. and Lind, J. M. (1980). Statistically based tests for the number of common factors. In *the annual meeting of the Psychometric Society. Iowa City, IA. 1980*.
16. Suhr, D. (2006). The basics of structural equation modeling. *Presented: Irvine, CA, SAS User Group of the Western Region of the United States (WUSS)*.
17. Suhr, D. (2008). Step your way through path analysis. *Western users of SAS software conference proceedings*.
18. Wright, S. (1921). Correlation and Causation. *Journal of agricultural research*, 20(7):557–585.
19. Yuan, K.-h., Chan, W., Marcoulides, G. A., and Bentler, P. M. (2015). Assessing Structural Equation Models by Equivalence Testing With Adjusted Fit Indexes. *Structural Equation Modeling: A Multidisciplinary Journal*, pages 1–12.