Comparison of Parameter Estimator Efficiency Levels of Path Analysis with Bootstrap and Jack Knife (Delete-5) Resampling Methods on Simulation Data

Adji Achmad Rinaldo Fernandes

Abstract

In practice, the assumptions of normality are often not met, this causes the estimation of the resulting parameters to be less efficient. Problems with the assumption that normality is not met can be overcome by resampling. The use of resampling allows data to be applied free from distribution assumptions. In this study, a simulation study was carried out by applying bootstrap resampling and jackknife resampling (delete-5) on path analysis assuming that the normality of the alignment was not met and the resampling amount set at 1000 with the degree of closeness between variables consisting of low closeness, medium closeness, high closeness and closeness level representing the level low to high closeness. Based on the simulation results, the resampling 1000 magnitude is able to overcome the problem of the assumptions of unmet normality. In addition, a comparison between bootstrap and jackknife resampling for conditions of side normality assumptions is not fulfilled and the closeness of the relationship between low, medium, high and closeness variables representing low to high closeness levels, the estimation results of path analysis parameters obtained by resampling jackknife are more efficient than resampling bootstrap.

Keywords: Resampling bootstrap, Resampling jackknife, and Efficiency.

1. INTRODUCTION

Path analysis is an analysis that can facilitate complex forms of relationships and the presence of intervening endogenous variables. Path analysis is the development of regression analysis. Path analysis has several assumptions, one of which is the residual normality or, in other words, the residual normal distribution. But, sometimes, the assumption of normality is not fulfilled, causing the resulting hypothesis testing less efficient. The problem of unfulfilled normality assumption can be overcome by resampling. The use of resampling allows data to be free from distribution assumptions or not to require normality assumption [11].

Based on previous studies, the Jackknife resampling method is more appropriate for estimating parameter values and parameter confidence interval values in the regression analysis with a narrow range of confidence intervals [1]. In algorithm studies of bootstrap and jackknife (delete-one) resampling methods for parameter estimation in the regression analysis, it is known that the parameter bias, standard errors, and confidence interval of Jackknife confidence are

Lecture, Brawijaya University, Department of Statistics, Faculty of Mathematics and Natural Sciences, Malang, East Java, Indonesia. E-mail: fernandes@staff.ub.ac.id

greater than the Bootstrap [8]. Therefore, in this research, we are interested in comparing the bootstrap and jackknife resampling methods in path analysis.

The data used in this research resulted from a simulation study. Based on the simulation study, Bootstrap resampling method can be used as an alternative method that can produce a very close regression parameter estimation to population parameters and fairly narrow confidence interval [12].

In contrast to previous studies, this research applied resampling methods in path analysis. Besides, this research was also intended to determine which resampling method is more efficient (bootstrap or jackknife (delete-5)) with a simulation study. Efficiency is relatively used to compare the variance of predictors between bootstrap and jackknife resampling methods.

2. LITERATURE REVIEW

Path analysis is an analysis developed from a linear regression analysis. According to [6], path analysis is an analysis used to evaluate direct and indirect effects through another cause. [9] stated that path analysis is a regression analysis with standardized variables. Another understanding suggests that path analysis determines the strength of the path shown in a path diagram [5]. Based on some of these notions, path analysis is an analysis used to study the direct and indirect effects of certain variables with standardized variables and it employs path diagrams to illustrate the relationships between variables.

According to [7], there are variable terms in path analysis. Exogenous variables are variables that are not determined by other variables in the model while endogenous variables are variables that are partly determined by other variables in the model.

2.1. Types of Effect on Path Analysis

According [9], there are three types of effect on path analysis. The three types of effect on path analysis are illustrated with the following path diagrams.

1. Direct Effect

Direct effects refer to the direct impact of exogenous variables on endogenous variables without intermediaries of other variables. Here is the path diagram illustrating direct effects.
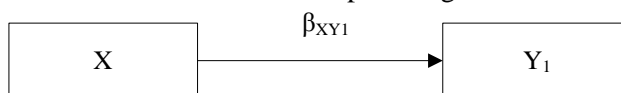
$$\beta_{XY1}$$

$$X \longrightarrow Y_1$$

Figure 2.1. Direct Effect

From Figure 2.1., it can be seen that the direct effect of exogenous variables on endogenous variables is β_{XY_1}.

2. Indirect Effect

Indirect effects refer to the impact of exogenous variables on endogenous variables that occurs through the intermediaries of other variables.
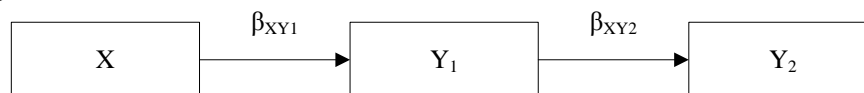
$$\beta_{XY1} \qquad \beta_{XY2}$$

$$X \longrightarrow Y_1 \longrightarrow Y_2$$

Figure 2.2. Indirect Effect

Based on Figure 2.2., variable X affects variable Y_2 through the intermediary of variable Y_1. The effect of variable X on variable Y_2 can not be measured directly but rather by multiplying the

direct effect of variable X on Variable Y_1 with the direct effect of variable Y_1 on Y_2. Mathematically, the indirect effect of Variable X on Y_2 is $\beta_{XY_1} \times \beta_{Y_1Y_2} = \beta_{XY_2}$.

   3. Total Effect
      The total effect is the total of all effects found in path analysis. The total effect can be measured by adding up all the direct and indirect effects.

2.2. Path Diagram
      Path diagram is a diagram used to illustrate the relationships between variables in path analysis. According to [2], path diagrams are useful for displaying patterns of causal relationships between variables graphically. In the path diagram, arrows between variables indicate causal directions, and the values written on the arrows represent effects in the form of regression coefficients. Here are some simple to more complex forms of path diagram according to [7].
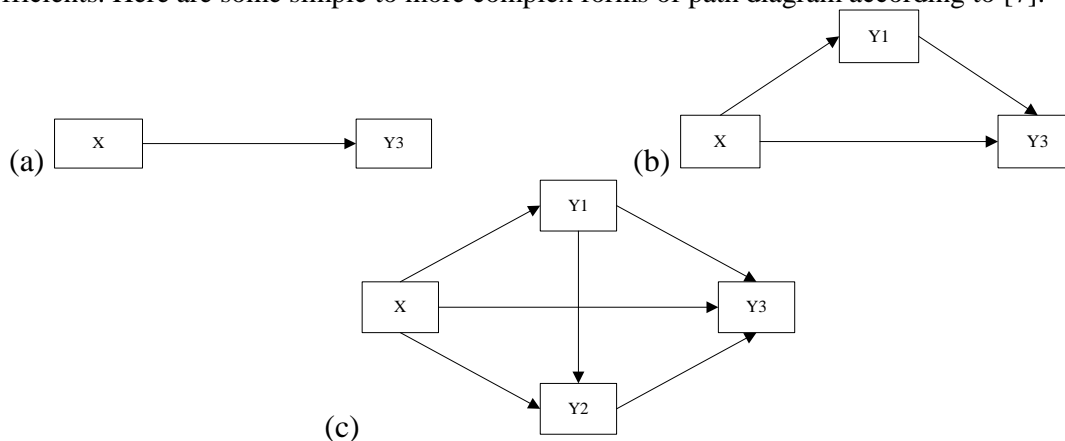

Figure 2.3. (a), (b), (c) Path Diagram

      Based on Figure 2.3. (c.), it can be seen that path diagrams illustrate the causal relationships of the four variables. Variable X is an endogenous variable while Variable Y_1, Y_2, and Y_3 are endogenous variables. Variable Y_3 is a pure endogenous variable while Variable Y_1 and Y_2 are intervening endogenous variables.

2.3. Path Analysis Model
      The path analysis in this research is following the path diagram as shown in Figure 2.3. (c). The path analysis model formed according to the diagram can be written as follows.

$$Y_{1i} = \beta_{XY_1}X_i + \varepsilon_{Y_1i} \tag{2.1}$$

$$Y_{2i} = \beta_{XY_2}X_i + \beta_{Y_1Y_2}Y_{1i} + \varepsilon_{Y_2i} \tag{2.2}$$

$$Y_{3i} = \beta_{XY_3}X_i + \beta_{Y_1Y_3}Y_{1i} + \beta_{Y_2Y_3}Y_{2i} + \varepsilon_{Y_3i} \tag{2.3}$$

      In Equation (2.1) to 2.3, i = 1, 2, ..., n, and n is the number of observations. In path analysis, the variables to be used are standardized first. The purpose of standardization is to equalize the averages and variances so that the path coefficient can be compared with other path coefficients [11]. Data standardization is done by standardizing the average to 0 and the variance to 1 with the following formula [6].

Adji Achmad Rinaldo Fernandes

$$Z_{X_i} = \frac{X_i - \bar{X}}{S} \tag{2.4}$$

Description:

Z_{X_i} : The value of variable X on i- standardized observation

X_i : The value of i-observation

$\bar{X}$: Average value

S : Standard deviation

After standardization, the form of path equation can be written as follows.

$$Z_{Y1i} = \beta_{Z_X Z_{Y_1}} Z_{X_i} + \varepsilon_{Z_{Y1i}}$$

$$Z_{Y2i} = \beta_{Z_X Z_{Y_2}} Z_{X_i} + \beta_{Z_{Y_1} Z_{Y_2}} Z_{Y1i} + \varepsilon_{Y2i} \tag{2.5}$$

$$Z_{Y3i} = \beta_{Z_X Z_{Y_3}} Z_{X_i} + \beta_{Z_{Y_1} Z_{Y_3}} Z_{Y1i} + \beta_{Z_{Y_2} Z_{Y_3}} Z_{Y2i} + \varepsilon_{Y3i}$$

In the form of a matrix, Equation (2.5) can be written like Equation (2.6).

$$
\begin{bmatrix} Z_{Y1i} \\ \vdots \\ Z_{Y1n} \\ Z_{Y2i} \\ \vdots \\ Z_{Y2n} \\ Z_{Y3i} \\ \vdots \\ Z_{Y3n} \end{bmatrix} =
\begin{bmatrix}
Z_{Xi} & 0 & 0 & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
Z_{Xn} & 0 & 0 & 0 & 0 & 0 \\
0 & Z_{Xi} & Z_{Y1i} & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
0 & Z_{Xn} & Z_{Y1n} & 0 & 0 & 0 \\
0 & 0 & 0 & Z_{X1i} & Z_{Y1i} & Z_{Y2i} \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & Z_{X1n} & Z_{Y1n} & Z_{Y2n}
\end{bmatrix}
\begin{bmatrix} \beta_{Z_X Z_{Y1}} \\ \beta_{Z_X Z_{Y2}} \\ \beta_{Z_{Y1} Z_{Y2}} \\ \beta_{Z_X Z_{Y3}} \\ \beta_{Z_{Y1} Z_{Y3}} \\ \beta_{Z_{Y2} Z_{Y3}} \end{bmatrix} +
\begin{bmatrix} \varepsilon_{Z_{Y1i}} \\ \vdots \\ \varepsilon_{Z_{Y1n}} \\ \varepsilon_{Z_{Y2i}} \\ \vdots \\ \varepsilon_{Z_{Y2n}} \\ \varepsilon_{Z_{Y3i}} \\ \vdots \\ \varepsilon_{Z_{Y3n}} \end{bmatrix} \tag{2.6}
$$

The above matrix form can be written into the following equation.

$$\underset{\sim}{Z}_Y = \mathbf{Z_X} \underset{\sim}{\beta} + \underset{\sim}{\varepsilon} \tag{2.7}$$

In the standardization process, the coefficient $\underset{\sim}{\beta}$ is equivalent to the correlation coefficient. Furthermore, estimation of path coefficients can be done based on path analysis model.

2.4. Path Coefficient Estimation

Estimation is done to obtain the path coefficient on the model. Path coefficients show the extent of influence between variables. One estimation of path analysis parameters is the Ordinary Least Square (OLS) method. The OLS method can only be used if linearity assumptions are met. Parameter estimation using the OLS method is done by minimizing the number of remaining squares.

$$\underset{\sim}{Y} = \mathbf{X} \underset{\sim}{\beta} + \underset{\sim}{\varepsilon} \tag{2.8}$$

From equation (2.8) we have:

$$\underset{\sim}{\varepsilon} = \underset{\sim}{Y} - \mathbf{X} \underset{\sim}{\beta} \tag{2.9}$$

Adji Achmad Rinaldo Fernandes

Furthermore, the remaining square can be written into $Q = \varepsilon^{\mathrm{T}} \varepsilon$. Thus, the OLS method minimizes the number of remaining squares as follows.

$$Q = \underset{\sim}{\varepsilon}^{\mathrm{T}} \underset{\sim}{\varepsilon}$$

$$= \left(\underset{\sim}{Y} - \mathbf{X}\underset{\sim}{\beta} \right)^{\mathrm{T}} \left(\underset{\sim}{Y} - \mathbf{X}\underset{\sim}{\beta} \right)$$

$$= \left(\underset{\sim}{Y}^{\mathrm{T}} - \underset{\sim}{\beta}^{\mathrm{T}}\mathbf{X}^{\mathrm{T}} \right) \left(\underset{\sim}{Y} - \mathbf{X}\underset{\sim}{\beta} \right)$$

$$= \underset{\sim}{Y}^{\mathrm{T}}\underset{\sim}{Y} - \underset{\sim}{Y}^{\mathrm{T}}\mathbf{X}\underset{\sim}{\beta} - \underset{\sim}{\beta}^{\mathrm{T}}\mathbf{X}^{\mathrm{T}}\underset{\sim}{Y} + \underset{\sim}{\beta}^{\mathrm{T}}\mathbf{X}^{\mathrm{T}}\mathbf{X}\underset{\sim}{\beta}$$

$$= \underset{\sim}{Y}^{\mathrm{T}}\underset{\sim}{Y} - 2\underset{\sim}{\beta}^{\mathrm{T}}\mathbf{X}^{\mathrm{T}}\underset{\sim}{Y} + \underset{\sim}{\beta}^{\mathrm{T}}\mathbf{X}^{\mathrm{T}}\mathbf{X}\underset{\sim}{\beta}$$

The equation of remaining squares above is derived to β and equates to zero as follows.

$$\frac{\partial(Q)}{\partial(\beta)} = 0$$

$$-2\mathbf{X}^{\mathrm{T}}\underset{\sim}{Y} + 2\mathbf{X}^{\mathrm{T}}\mathbf{X}\underset{\sim}{\widehat{\beta}} = 0$$

$$-\mathbf{X}^{\mathrm{T}}\underset{\sim}{Y} + \mathbf{X}^{\mathrm{T}}\mathbf{X}\underset{\sim}{\widehat{\beta}} = 0$$

$$\mathbf{X}^{\mathrm{T}}\mathbf{X}\underset{\sim}{\widehat{\beta}} = \mathbf{X}^{\mathrm{T}}\underset{\sim}{Y}$$

$$\underset{\sim}{\widehat{\beta}} = \left(\mathbf{X}^{\mathrm{T}}\mathbf{X} \right)^{-1} \mathbf{X}^{\mathrm{T}}\underset{\sim}{Y} \tag{2.10}$$

2.5. Path Analysis Asumption
According to [4] and [10], assumptions underlying path analysis are as follows:
1. Linearity Assumptions
   The path analysis model assumes that the relationship between variables is linear. According to [4], linearity assumptions can be tested using the Ramsey's Regression Specification Error (RESET) test. Here are the steps for Ramsey's RESET.
   a. For example, the first equation is a linear r ssion model $Y_i = \alpha_0 + \alpha_1 x_i + \varepsilon_i$. Then, parameter estimation from the equation is conducted. After that, the determination coefficient is calculated with the following formula (2.10).

$$R_1^2 = 1 - \frac{\sum_{i=1}^{n} \left(Y_i - \widehat{Y}_i \right)^2}{\sum_{i=1}^{n} \left(Y_i - \overline{Y} \right)} \tag{2.10}$$

   b. For example, the second equation is a quadratic model, $Y_i = \alpha_0 + \alpha_1 x_i + \alpha_2 \widehat{Y}_i^2 + \varepsilon_i$. Similar to the previous stage, parameter estimation is then also made in a quadratic model. Moreover, the determination coefficient is calculated with the following formula (2.11).

$$R_2^2 = 1 - \frac{\sum_{i=1}^{n} \left(Y_i - \hat{Y}_i\right)^2}{\sum_{i=1}^{n} \left(Y_i - \bar{Y}\right)} \tag{2.11}$$

c. Hypothesis testing for Ramsey's RESET test is conducted with test statistics following F distribution.

$H_0 : \alpha_2 = 0$

$H_1 : \alpha_2 \neq 0$

Test Statistics:

$$F = \frac{\left(R_2^2 - R_1^2\right)/2}{\left(1 - R_2^2\right)\left(n - \left(k + 2\right)\right)} \; \Box \; F_{(k-1, n-k-2)} \tag{2.12}$$

2. The pattern of relationships between variables is recursive. It means that there is only a one-way causal relationship in the model, or in other words, there are no endogenous variables that have a reciprocal effect. Endogenous variables in path analysis are at least measured using interval scales.

2.5. Central Limit Theorem

The central limit theorem has an important role in the sampling distribution. One of its roles is to estimate the parameters of the average and variance of a population. According to Mandenhall (1981), supposing that $Y_1, Y_2, ..., Y_n$ are free and spread randomly and identically with $E(Y_i) = \mu$, $V(Y_i) = \sigma^2 < \infty$ and, then the distribution function of U_n will converge to a standard normal distribution where n approaches infinity. U_n can be defined as the following equation.

$$U_n = \sqrt{n}\left(\frac{\bar{Y} - \mu}{\sigma}\right) \tag{2.13}$$

The function of the random variable U_n with Equation (2.14) can be approximated by the following equation.

$$P(\text{a} \leq \text{U}_n \leq b) \rightarrow \int_a^b \left(\frac{a}{\sqrt{2\pi}}\right) e^{-u^2/2} du \tag{2.14}$$

If U_n (n-sized random variable) with $n > 30$, it can be approximated by the standard normal distribution.

2.6. Resampling

Resampling is the process of drawing repeated samples from existing or original samples so that a new sample is obtained. The new sample is obtained from original sized samples taken at random either with or without returns. The re-sampling method can be applied as an alternative if the number of observations does not meet the research need which can cause the parameter estimation incorrect. Besides, the implementation of the resampling method allows the validity of data that is free from assumptions or, in other words, does not need the normality assumption [11].

Adji Achmad Rinaldo Fernandes

2.6.1. Bootstrap

Bootstrap was first introduced by Efron in 1979. Bootstrap is a resampling method for estimating the parameters of each sample of bootstrap B which is n-sized random samples and taken with returns. Resampling is done for B times in which the number of B required is large enough so that the estimated parameters obtained is convergent. Supposing $\hat{F}$ of an empirical distribution with selected observation opportunities becomes a bootstrap sample of $1/n$ for each value of observation x_i, then $i = 1, 2, ..., n$.

[3] defined bootstrap samples as n-sized random samples taken from $\hat{F}$. Bootstrap samples can be written as $\underset{\sim}{x}^* = \left(x_1^*, x_2^*, ..., x_n^* \right)$ or can be written as Equation (2.15) below.

$$\hat{F} \to \underset{\sim}{x}^* = (x_1^*, x_2^*, ..., x_n^*) \tag{2.15}$$

Star notation indicates that $\underset{\sim}{x}^*$ is the result of a random process or resampling from a set of original data. Meanwhile, $x_1^*, x_2^*, ..., x_n^*$ are observations in the bootstrap samples obtained randomly with the return of the original data set $x_1, x_2, ..., x_n$.

From equation (2.15), $\underset{\sim}{x}_1^*, \underset{\sim}{x}_2^*, ..., \underset{\sim}{x}_n^*$ are paired data selected to be bootstrap samples. Based on the explanation above, the bootstrap resampling process, in general, can be depicted as follows Figure 2.4.
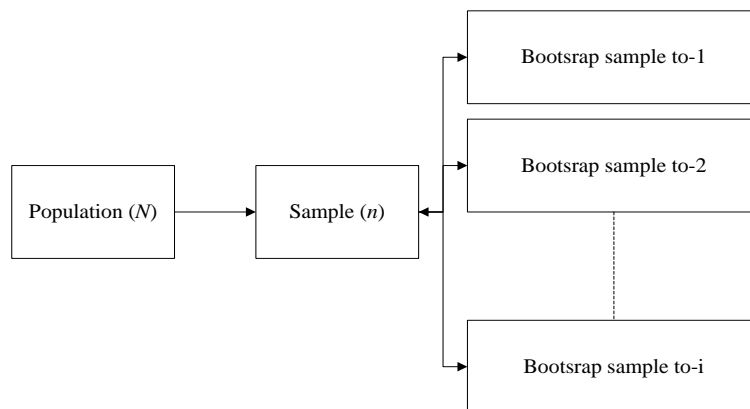

Figure 2.4. Bootstrap Resampling Process

The steps of the bootstrap method for estimating standard errors are as follows [3].

1. Determining the number of B as the size of resampling to obtain bootstrap samples $\mathbf{x}^{*1}, \mathbf{x}^{*2}, ..., \mathbf{x}^{*B}$ from random sampling with a return of n elements from .the original data sample.

2. Calculating bootstrap replication for each bootstrap sample. The function applied in each bootstrap sample function is the same as the one in the original sample. If the parameter estimation in the original sample is done with OLS, the parameter estimation in the bootstrap sample will also be conducted with OLS

$$\hat{\beta}^*(b) = s\left(\mathbf{x}^{*b} \right), \text{ where } b = 1, 2, ..., B$$

$s\left(\mathbf{x}^{*b} \right)$ is the result of the application of the same function in the original data sample set and the bootstrap data sample set.

3. Estimating standard errors by using standard deviations for bootstraps replicated as many as *B* times.

$$\widehat{se}_B = \sqrt{\frac{\sum\limits_{b=1}^{B}\left[\widehat{\beta}*(b)-\widehat{\beta}*(.)\right]}{(B-1)}} \qquad (2.16)$$

where:

$$\widehat{\beta}*(.)=\sum\limits_{b=1}^{B}\frac{\widehat{\beta}*(b)}{B} \qquad (2.17)$$

Description:

$\widehat{se}_B$: bootstrap standard error

$\widehat{\beta}*(b)$: parameter estimator on the b-bootstrap process

$\widehat{\beta}*(.)$: average parameter estimator of the bootstrap process

B : the size of resampling

2.6.2. Jackknife

Jackknife is a resampling method introduced by Quenouille in 1949 for estimation of bias. Then, in 1958, Tukey introduced jackknife to estimate standard deviations. The principle of the jackknife method is to eliminate five observations from *n*-sized samples and take other observations without returns. In the next stage, the deleted sample is returned and so are the other five observations until all observations from the population have a chance to be deleted. Based on the process, the Jackknife resampling process generally can be seen as shown in the following Figure 2.5.
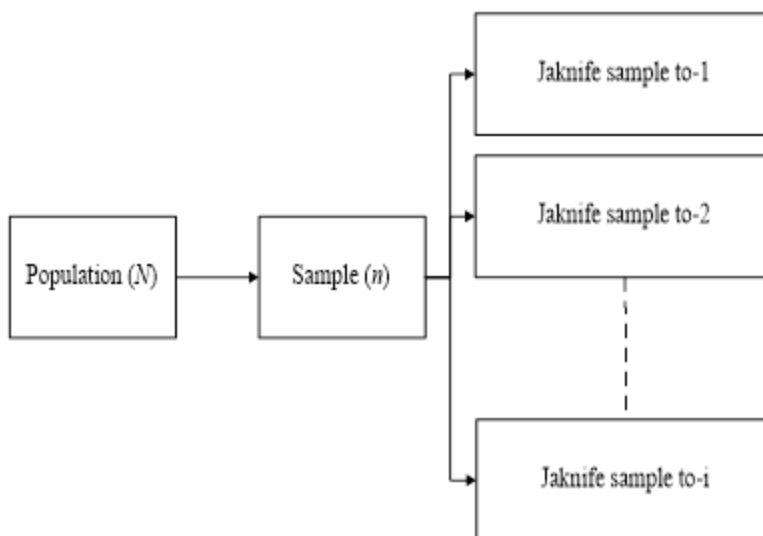

Figure 2.5. Jackknife Resampling Process

Based on the process, it is obtained that the $n-d$ - sized sample where *n* is the initial sample and *d* is the number of observations deleted in the Jackknife sampling process. Therefore, the

number of Jackknife samples that may be formed is $\begin{pmatrix} n \\ d \end{pmatrix}$ samples. The steps for the Jacknife

resampling method are explained as follows.
1. Conducting resampling by removing five observation lines alternately from the original sample set on each Jacknife sample.
2. Calculating the replication of i-Jackknife from $\hat{\beta}$.

$$\hat{\beta}_{(i)} = s\left(x_{(i)}\right)$$

3. Calculating the standard error for the Jacknife sample.

$$\hat{se}_{jk} = \sqrt{\frac{n}{d\begin{pmatrix} n \\ d \end{pmatrix}} \sum_{i=1}^{n} \left(\hat{\beta}_{(i)} - \hat{\beta}_{(.)}\right)^2} \tag{2.18}$$

Description:

$\hat{se}_{jk}$: Jackknife standard error

$\hat{\beta}_{(i)}$: Parameter estimator on the i-Jackknife process

$\hat{\beta}_{(.)}$: Average parameter estimator on the i-Jackknife process

2.7. Relative Efficiency

Comparison of the resampling methods is measured based on the relative efficiency value. According to [13], relative efficiency is calculated by comparing the variance between the two-parameter estimators. The relative efficiency of the two estimators can be written as follows.

$$eff\left(\hat{\beta}_{JK}, \hat{\beta}_{BS}\right) = \frac{V\left(\hat{\beta}_{BS}\right)}{V\left(\hat{\beta}_{JK}\right)} \tag{2.19}$$

Description:

$eff(\hat{\beta}_{BS}, \hat{\beta}_{JK})$: efficiency between bootstrap resampling method estimators and jackknife resampling method estimators

$V\left(\hat{\beta}_{JK}\right)$: variance of parameter estimators by the jackknife resampling method

$V\left(\hat{\beta}_{BS}\right)$: variance of parameter estimators by the bootstrap resampling method

Comparison of efficiency between Bootstrap and Jacknife resampling estimator variance. The variance of parameter estimators with Jacknife resampling method. If the calculated efficiency result using Equation (2.19) is more than 1, Estimator $\hat{\beta}_{JK}$ is said more efficient than Estimator $\hat{\beta}_{BS}$. Conversely, if the calculated efficiency result is less than 1, Estimator $\hat{\beta}_{BS}$ is said more

efficient than Estimator $\widehat{\beta}_{JK}$. Meanwhile, if the calculate efficiency result is equal to 1, both estimators are equally efficient.

3. METHODOLOGY

The data used in this research are simulation study data with one exogenous variable, two intervening endogenous variables, and one pure endogenous variable. The exogenous variable was raised and standardized while the endogenous variables were determined based on the values of exogenous variables, path coefficients, and residuals generated following the Weibull distribution. The path coefficients in this research were raised at several levels of closeness with a value range of $|0,05|-|0,15|$ representing a low closeness of relationships, a value range of $|0,20|-|0,40|$ representing a medium closeness of relationships, a value range of $|0,50|-|0,90|$ representing a high closeness of relationships, and $|0,05|-|0,90|$ representing the three levels of closeness.

4. RESULT AND DISCUSSION

The relative efficiency results for the low level of closeness can be seen in Table 4.1.

Table 4.1. Relative Efficiency of Parameter Estimators at Low Level of Closeness

Path Coefficients	Ragam *Bootstrap*	Ragam Jackknife	Relative Efficiency
β_{XY1}	0.00007	0.00002	3.73064
β_{XY2}	0.00017	0.00004	3.65072
β_{Y1Y2}	0.01402	0.00395	3.62942
B_{XY3}	0.00044	0.00012	3.70288
β_{Y1Y3}	0.03725	0.01065	3.70772
β_{Y2Y3}	0.03932	0.01139	3.71637

From Table 4.1., it can be seen that the relative efficiency values of all path analysis coefficients were > 1, indicating that the Jackknife method had a smaller variance than the Bootstrap method. Therefore, it can be concluded that path analysis with the Jacknife resampling method is three times more efficient than that with the bootstrap resampling method.

The relative efficiency results for the medium level of closeness can be seen in Table 4.2.

Table 4.2. Relative Efficiency of Parameter Estimators at Medium Level of Closeness

Path Coefficients	Ragam *Bootstrap*	Ragam *Jackknife*	Relative Efficiency
β_{XY1}	0.00007	0.00002	3.7306
β_{XY2}	0.00128	0.00036	3.6299
β_{Y1Y2}	0.01402	0.00395	3.6294
B_{XY3}	0.00360	0.00100	3.7194
β_{Y1Y3}	0.03586	0.01018	3.7215
β_{Y2Y3}	0.03932	001139	3.7163

Adji Achmad Rinaldo Fernandes

From Table 4.2., it can be seen that the relative efficiency values of all path analysis coefficients were > 1, indicating that the Jackknife method had a smaller variance than the bootstrap method. Therefore, it can be concluded that path analysis with the Bootstrap resampling method.

The relative efficiency results for the high level of closeness can be seen in Table 4.3.

Table 4.3. Relative Efficiency of Parameter Estimators at High Level of Closeness

Path Coefficients	Ragam *Bootstrap*	Ragam *Jackknife*	Relative Efficiency
β_{XY1}	0.00007	0.00019	3.73064
β_{XY2}	0.00683	0.00192	3.62889
β_{Y1Y2}	0.01402	0.00394	3.62942
B_{XY3}	0.01937	0.00536	3.72607
β_{Y1Y3}	0.04226	0.01193	3.72165
β_{Y2Y3}	0.03932	0.01139	3.71637

From Table 4.3., it can be seen that the relative efficiency values of all path analysis coefficients were > 1, indicating that the Jackknife method had a smaller variance than the Bootstrap method. Therefore, it can be concluded that path analysis with the Jacknife resampling method is three times more efficient than that with the bootstrap resampling method.

The relative efficiency results for the low-to-high level of closeness can be seen in Table 4.4.

Table 4.4. Relative Efficiency of Parameter Estimators at Low to High Level of Closeness

Path Coefficients	Ragam *Bootstrap*	Ragam *Jackknife*	Relative Efficiency
β_{XY1}	0.00007	0.00002	3.7306
β_{XY2}	0.00128	0.00036	3.6229
β_{Y1Y2}	0.01402	0.00394	3.6294
B_{XY3}	0.00360	0.00100	3.7267
β_{Y1Y3}	0.03868	0.01018	3.7039

From Table 4.4., it can be seen that the relative efficiency values of all path analysis coefficients were > 1, indicating that the Jackknife method had a smaller variance than the Bootstrap method. Therefore, it can be concluded that path analysis with the Jacknife resampling method is three times more efficient than that with the bootstrap resampling method. In contrast to [8] study, the Jacknife resampling in this research was done by removing one observation only so that the resulting Jacknife samples were less than the bootstrap samples. This caused Bootstrap resampling to be better than Jacknife resampling.

5. CONCLUSION

Based on the simulation study that has been done, the use of Bootstrap and Jackknife resampling methods on data with normality assumption is not fulfilled, indicating that both bootstrap and jackknife resampling methods can be applied and able to overcome normality assumption. The calculated relative efficiency results in various closeness levels of relationship show that the Jacknife resampling method (delete-5) is three times more efficient than the Bootstrap resampling method.

REFERENCES

[1] Ariani, D., Yuki N. N., Desi. 2017. *Perbandingan Metode Bootstrap dan Jackknife Resampling dalam Menentukan Nilai Estimasi dan Interval Kontingensi Parameter Regresi, Jurnal Eksponensial*. Vol 8 Issue 1, 43-49.

[2] Dillon, W. R. dan Goldstein, M. 1984. *Multivariate Analysis Methods and Application*. New York: John Wiley & Sons Inc.

[3] Efron, B. dan Tibhirani, R.J. 1993. *An Intoduction to the Bootstrap*. US: Springer.

[4] Gujarati, D. 2004. *Basic Econometrics*. Fourth Edition. New York: McGraw Hill.

[5] Rutherford, R.D. 1993. *Statistikal Model For Causal Analysis*. New York: John Wiley & Sons Inc.

[6] Hair, J. F., Black, W. C., Babin, B. J., dan Anderson, R. E. (1998). *Multivariate Data Analysis*. New Jersey: Prentice Hall.

[7] Li, C. C. 1975. *Path Analysis* – A Primer. USA: The Boxwood Press.

[8] Sahinler S, dan Topuz D. 2007. *Bootstrap and Jackknife Resampling Algorithm for Estimation of Regression Parameters*. Journal of Applied Quantitative Methods.

[9] Solimun. 2010. *Analisis Multivariat Pemodelan Struktural*. Malang: CV Citra Malang.

[10] Solimun, 2002. *Multivariate Analysis Structural Equation Modelling (SEM) Lisrel dan Amos*. Malang: Penerbit Universitas Negeri Malang.

[11] Solimun, A. A. R. Fernandes, dan Nurjannah. 2017. *Metode Statistika Multivariat: Pemodelan Persamaan structural (SEM) perndekatan WarpPLS*. Malang: UB Press.

[12] Sungkono, J. (2015). Bootstrap Resampling Observasi Pada Estimasi Parameter Regresi Menggunakan Software R. *MAGISTRA*, *27*(92).

[13] Wackerly, D.D., Mendenhall, W., dan Scheaffer, R.L, 1981. *Mathematical Statistics with Application Second Edition*. Boston, Mass: Duxbury Press.