BAREKENG: Journal of Mathematics and Its Applications
June 2023 Volume 17 Issue 2 Page 0807–0818
P-ISSN: 1978-7227 E-ISSN: 2615-3017
https://doi.org/10.30598/barekengvol17iss2pp0807-0818

COMPARISON OF RESAMPLING EFFICIENCY LEVELS OF JACKKNIFE AND DOUBLE JACKKNIFE IN PATH ANALYSIS

M. Fikar Papalia[1*], Solimun[2], Nurjannah[3]

[1,2,3,] *Department Statistics, Faculty Mathematics and Science, Brawijaya University, Veteran Street, Ketawanggede, Lowokwaru, Malang City, 65145, Indonesia*

Corresponding author's email: [*] ***m.fikarpapalia@gmail.com***

Article History:

Received: 8th December 2022
Revised: 12th April 2023
Accepted: 15th April 2023

Keywords:

Double Jackknife Resampling;
Efficiency;
Jackknife Resampling.

ABSTRACT

The assumption of normality is often not fulfilled. This causes the estimation of the resulting parameters to be less efficient. The problem with assuming that normality is not satisfied can be overcome by resampling. The use of resampling allows data to be applied free of distributional assumptions. In this study, a research simulation was carried out by applying Jackknife and Double Jackknife resampling in path analysis with the assumption that the normality of the residuals was not fulfilled and the number of resampling was set at 100 with the degree of closeness level of relationship between variables consisting of low, medium, and high closeness. Based on the simulation results, resampling with a power of 100 can overcome the problem of unfulfilled normality assumptions. In addition, the comparison of the relative efficiency level of the Jackknife and Double Jackknife resampling in the path analysis obtained that the Double Jackknife resampling has more efficiency than the Jackknife resampling.

This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution-ShareAlike 4.0 International License.

How to cite this article:

M. F. Papalia and S. Nurjannah., "COMPARISON OF RESAMPLING EFFICIENCY LEVELS OF JACKKNIFE AND DOUBLE JACKKNIFE IN PATH ANALYSIS," *BAREKENG: J. Math. & App.*, vol. 17, iss. 2, pp. 0807-0818, June, 2023.

Copyright © 2023 Author(s)
Journal homepage: *https://ojs3.unpatti.ac.id/index.php/barekeng/*
Journal e-mail: *barekeng.math@yahoo.com; barekeng.journal@mail.unpatti.ac.id*

Research Article • Open Access

1. INTRODUCTION

Path analysis is an extension of multiple linear regression analysis, which has more than one equation in the form of a system. In path analysis, the terms response and predictor variables are no longer used. Instead, the terms exogenous and endogenous variables are used **[1]**. As an extension of multiple linear regression analysis, the assumptions in estimating the path coefficient are almost the same as the assumptions that apply in multiple linear regression analysis. The assumptions that must be met in the path analysis are the assumption of linearity, residual variance homoscedasticity, and residual normality assumption **[2]**. The assumption of normality of the residuals in path analysis is as important as it is in regression analysis. There are times when the residual normality assumption in the path analysis cannot be fulfilled, so it needs to be handled.

If the residual normality assumption cannot be met, then several things can be done, for example, transforming the data, trimming the outlier data, or adding observations **[3]**. In addition to the method already mentioned, there is one more method that can be used to overcome violations of the residual normality assumption, namely *resampling* **[4]**.

Resampling is a method of taking repeated samples from the same sample **[5]**. Bootstrap and Jackknife methods are nonparametric and resampling techniques that aim to estimate standard errors and bias values, Jackknife itself is an alternative to bootstrap **[6]**. **[7]** Bootstrap and Jackknife are two methods used to estimate an unknown population distribution with empirical distributions obtained from the resampling process. The resampling used in this study is Jackknife and Double Jackknife resampling.

The basic principle of the Jackknife method is the removal of the first element from the original data **[8]**, the result of deleting the first row is called the first stage Jackknife data as much as β_1 resampled again as much as β_2 replication so that it is called the second stage double Jackknife. The weakness of the Double Jackknife method is that it takes longer to calculate because it has to calculate as many $\beta_1 + \beta_1\beta_2$ test statistic value.

The 100 data used in this research is a simulation study using generated data with one exogenous variable, one intervening endogenous variable, and one pure endogenous variable **[9]**. These three variables are measured directly (observable variables), so they do not require a measurement model. Thus, the data has an interval or ratio scale. Exogenous variables are determined by standardizing $\bar{x} \pm 2s$, where $\bar{x} = 0$ and $s = 1$. The distance between observations on exogenous variables is made the same. Pure endogenous variables are calculated through a linear regression function with three variations of the path coefficient **[10]**. Path coefficients with a range of 0.05 – 0.20 describe a low closeness relationship, a value range of 0.20 – 0.50 describe a medium closeness relationship, and a value range of 0.50 – 1.00 describes a high closeness relationship.

The comparison of the Jackknife and Double Jackknife methods is said to be good in terms of relative efficiency. Comparative research on the efficiency of the parameter estimator from path analysis by applying Jackknife (delete-5), the Jackknife resampling method is more appropriate for parameter estimation in path analysis, this is indicated by a relative efficiency above one [6]. The use of path analysis with the jackknife method proved to be more effective, and resulted in relatively good asymptotic assumptions [11].

In this study, the effectiveness was tested by applying the Double Jackknife resampling method, where the process will resampling the first stage of resampling data. The novelty of this research is the use of the Double Jackknife method applied to simulated data.

In this study, residuals from generated data are used, namely residuals that follow an exponential distribution [12]. The exponential distribution represents the condition when the residual normality assumption is not met. Comparison of the Efficiency Level of the Parameter Estimator from Path Analysis with Bootstrap and Jack Knife (Delete-5) with Simulation Data shows that the jackknife resampling method (delete-5) is three times more efficient than the bootstrap resampling method **[6]**. This study applies the resampling method in path analysis. In addition, this research is also intended to determine which resampling method is more efficient, Jackknife or Double Jackknife with a simulation study.

2. RESEARCH METHODS

2.1 Path Analysis

Path analysis is used when research analyses relationships between complex variables that cannot be done using multiple regression. In complex relationships or more than one dependent variable, a series of regression equations is needed. Path analysis was developed as a method to study the direct and indirect effects of the independent variables on the dependent variable [13].

Path analysis is a technique for analysing cause-and-effect relationships that occur in multiple regression if the independent variables affect the dependent variable not only directly but also indirectly. Another definition says path analysis is the direct development of multiple regression forms to provide estimates of the magnitude and significance of a hypothetical causal relationship in a set of variables. Another definition says path analysis is an extended regression model that is used to test the alignment of the correlation matrix with two or more causal relationship models desired by the researcher [14].

This analysis is a method for explaining and looking for causal relationships between variables. Path analysis is used to examine the relationship between causal models that have been formulated by researchers based on theoretical considerations and certain knowledge. Besides being based on data, causal relationships are also based on knowledge, hypothesis formulation, and logical analysis. This path analysis can be used to test a set of causal hypotheses as well as interpret those relationships [15]. Based on the description above, path analysis is not a method to find causes but an applied method for causal models formulated by researchers with a knowledge base and theoretical considerations.

2.2 Types of Influence on Path Analysis

A variable can be viewed as a cause or an effect. This can be seen from the influence of the variables [16]. It explained that there are three types of influence in the path analysis, namely:

1) Direct Effect
   Inter-variables are said to directly influence when the influence between exogenous and endogenous variables occurs without going through other variables as intermediaries.

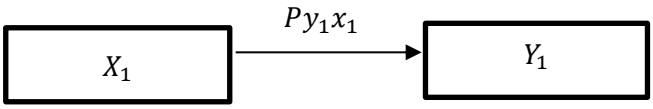

Figure 1. Direct influence

From Figure 1 it can be seen that the magnitude of the direct effect can be known directly as Py_1x_1.

2) Indirect Effect
   Inter-variables have an indirect effect if the influence between exogenous and endogenous variables is through other variables as intermediaries.

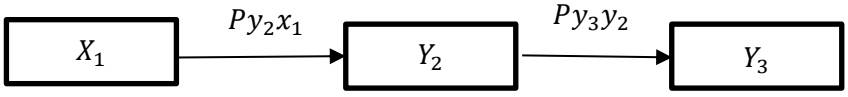

Figure 2. Indirect effect

From Figure 2, it can be seen that the effect of X_1 is Y_3 through Y_2. The amount of direct influence is calculated by diverting the direct influence of X_1 to Y_2 and the direct influence of Y_2 to Y_3 with the formula $Py_2x_1 \times Py_3y_2$.

3) Total Effect
   The total effect is the sum of the step effect and the indirect effect. Based on Figure 2, the total effect can be calculated using the formula $Py_1x_1 + Py_2x_1 \times Py_3y_2$.

2.3 Path Charts

Path diagrams are an important component of path analysis. Through the path diagram, it can be seen the direct and indirect influence on the relationship between exogenous and endogenous variables. Cause and effect relationships in the path diagram are shown in the direction of the arrows. In path analysis, there is at least one exogenous variable (X), mediating (Y_1), and endogenous (Y_2). The shape of the path diagram with three variables is as follows.

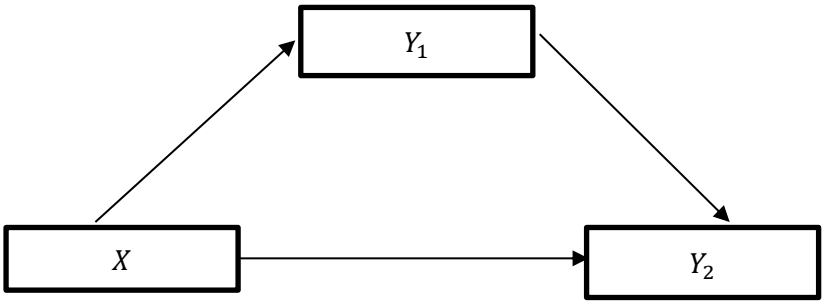

Figure 3. Path diagram with three variables

Figure 3 shows that X affects Y_1 and Y_2, Y_1 affects Y_2 , while Y_2 is influenced by X and Y_1.

Before creating a model, it is necessary to standardize the variables in the path analysis so that they have the same mean and variance. Thus, the path coefficients obtained usually have the same units and can be compared. The following is a form of standardization carried out on exogenous variables.

$$Z_{X_i} = \frac{X_i - \bar{X}}{S}; with\ S = \sqrt{\frac{(X_i - \bar{X})^2}{n-1}} \quad (1)$$

By carrying out standardization, each variable will follow the standard normal distribution, namely the normal distribution with $mean = 0$ and $variance = 1$. To get the original observed value, the reverse transformation can be carried out as follows.

$$\begin{aligned} Z_X &= \frac{X_i - \bar{X}}{S} \\ S \times Z_X &= X_i - \bar{X} \\ X_i &= (S \times S_X) + \bar{X} \end{aligned} \quad (2)$$

The path analysis model is a system of equations. This model can be formed based on the path diagram. It is necessary to solve the system of equations simultaneously (simultaneously), starting from parameter estimation and hypothesis testing to interpretation.

The system of equations obtained from the path diagram in **Figure 3** is as follows.

$$Y_{1i} = \beta_{XY_1} X_i + \varepsilon_{Y_1 i} \quad (3)$$
$$Y_{2i} = \beta_{XY_2} X_i + \beta_{Y_1 Y_2} Y_{1i} + \varepsilon_{Y_2 i} \quad (4)$$

In the equation above, i moves from $1,2,\cdots,n$, where n indicates the number of observations. After being standardized with **Equation (1)**, the system of **Equations (3-4)** becomes as follows.

$$\begin{aligned} Z_{Y_1 i} &= \beta_{Z_X Z_{Y_1}} Z_{Xi} + \varepsilon_{Z_{Y_1}} \\ Z_{Y_2 i} &= \beta_{Z_X Z_{Y_2}} Z_X + \beta_{Z_{Y_1} Z_{Y_2}} Z_{Y_1 i} + \varepsilon_{Z_{Y_2 i}} \end{aligned} \quad (5)$$

In matrix form, the system of **Equation (5)** can be written as follows.

$$\begin{pmatrix} Z_{Y_{11}} \\ \vdots \\ Z_{Yn} \\ Z_{Y_{21}} \\ \vdots \\ Z_{Y_{2n}} \end{pmatrix} = \begin{pmatrix} Z_{X_1} & 0 & 0 \\ \vdots & \vdots & \vdots \\ Z_{X_n} & 0 & 0 \\ 0 & Z_{X_1} & Z_{Y_{11}} \\ \vdots & \vdots & \vdots \\ 0 & Z_{X_n} & Z_{Y_{1n}} \end{pmatrix} \begin{pmatrix} \beta_{Z_X Z_{Y_1}} \\ \beta_{Z_X Z_{Y_2}} \\ \beta_{Z_{Y_1} Z_{Y_2}} \end{pmatrix} + \begin{pmatrix} \varepsilon_{Z_{Y_{11}}} \\ \vdots \\ \varepsilon_{Z_{Y_{1n}}} \\ \varepsilon_{Z_{Y_{21}}} \\ \vdots \\ \varepsilon_{Z_{Y_{2n}}} \end{pmatrix} \quad (6)$$

$$\underset{\sim}{Y} = \boldsymbol{X}\underset{\sim}{\beta} + \underset{\sim}{\varepsilon}$$

2.4 Path Coefficient Estimation

The path coefficient shows the magnitude of the direct influence of exogenous variables on endogenous variables in a system of equations. One method that can be used to estimate the path coefficient is Ordinary Least Square (OLS).

The principle of the OLS method is to minimize the sum of the residual squares. Based on **Equation (6)**, $Y = \mathrm{X}\beta + \varepsilon$ can be written as $\varepsilon = Y - \mathrm{X}\beta$. Thus, the sum of the squares of the residuals can be written as $Q = \varepsilon^T \varepsilon$.

The OLS method minimizes the following functions.

$$\min\{Q\} = \min\{\varepsilon^T \varepsilon\} = \min\{(Y - X\beta)^T (Y - X\beta)\}$$
$$\min\{Q\} = \min\{\varepsilon^T \varepsilon\} = \min\{(Y - X\beta)^T (Y - X\beta)\} \quad (7)$$

2.5 Resampling

Resampling is the process of repeating sampling from an existing or original sample so that a new sample is obtained **[17]**. A new sample is obtained from the original sized sample taken at random, either with replacement or without replacement. The resampling method can be applied as an alternative if the number of observations does not meet the needs of the research, which can lead to inaccurate parameter estimates **[18]**. In addition, the application of the resampling method allows the validity of the data, which is free from assumptions or, in other words, does not require normality assumptions.

2.5.1 Jackknife

Jackknife is a resampling method introduced by Quenouille to estimate bias. Tukey introduced Jackknife to estimate the standard deviation **[19]**. The Jackknife method is used for taking new samples repeatedly from the original data of size n by deleting the $ke - i$ observation with $i = 1,2,3, \ldots, n$. then the Jackknife simulation is based on the new data set $\times^* = x_1^*, x_2^*, \cdots, x_n^*$ which is used to take new samples repeatedly from the original data of size n by removing the $i - th$ observation.
Jackknife's resampling process in general can be seen as shown below in **Figure 4**.

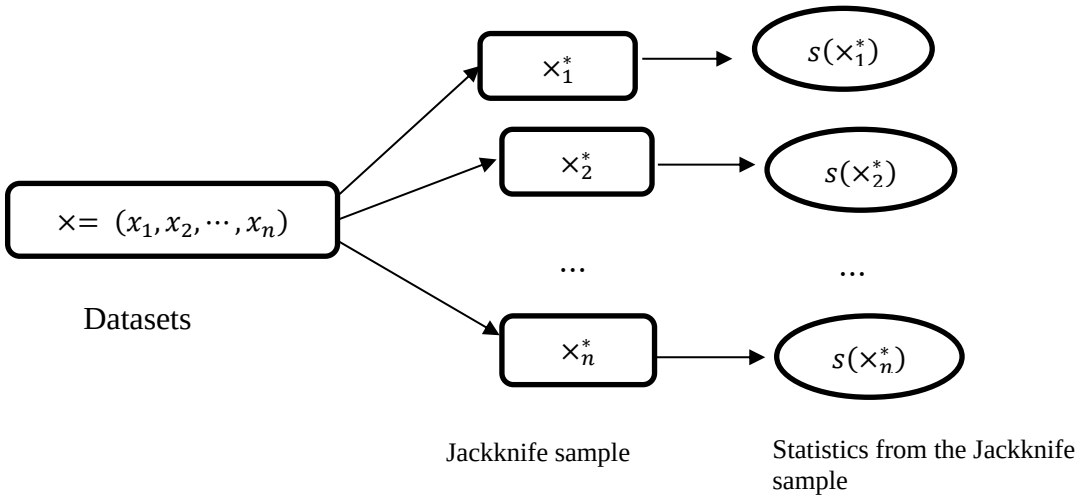

Figure 4. Schematic of the Jackknife process

Based on this process, the collection of observational data (observation) is based on the removal of one sample or group of samples from the initial sample, which is considered as a population. At one stage and the next, the removed sample is returned and one or a group of samples is deleted and so on, until all samples from the population have had a chance to be deleted.

The principle is to remove one piece of data and repeat it as many times as there are samples. To estimate the regression parameters using the Jackknife procedure, eliminating one piece of data can be done using the following procedure.

Take a random sample of size n, where:

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} and\ X = \begin{bmatrix} 1 & x_{11} & x_{12} & \dots & x_{1j} \\ 1 & x_{21} & x_{22} & \dots & x_{2j} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \dots & x_{nj} \end{bmatrix} \text{ is the actual sample.}$$

The next step in jackknife is to remove one row from the vector, for the Jackknife to remove the first line in the vector as follows:

$$Y^{**1} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} and\ X^{**1} = \begin{bmatrix} 1 & x_{21} & x_{22} & \dots & x_{2j} \\ 1 & x_{32} & x_{32} & \dots & x_{3j} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \dots & x_{nj} \end{bmatrix} \tag{8}$$

The data that has been omitted from the first row in the vector is called the Jackknife data and is solidly denoted in the following form:

$$Y^{**i} = \begin{bmatrix} y_1^{**i} \\ y_2^{**i} \\ \vdots \\ y_n^{**i} \end{bmatrix};\ X^{**1} = \begin{bmatrix} 1 & x_{11}^{**i} & x_{12}^{**i} & \dots & x_{1j}^{**i} \\ 1 & x_{21}^{**i} & x_{22}^{**i} & \dots & x_{2j}^{**i} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{(n-1)1}^{**i} & x_{(n-1)2}^{**i} & \dots & x_{(n-1)j}^{**i} \end{bmatrix};\ \varepsilon^{**i} = \begin{bmatrix} e_1^{**i} \\ e_2^{**i} \\ \vdots \\ e_n^{**i} \end{bmatrix} \tag{9}$$

To estimate the parameter $\hat{\beta}^{**i}$ search using the least squares method to minimize the sum of squared errors as follows:

$$\begin{aligned} e^{**i^T} e^{**i} &= \left(Y^{**i} - X^{**i}\hat{\beta}^{**i}\right)^T \left(Y^{**i} - X^{**i}\hat{\beta}^{**i}\right) \\ &= \left(Y^{**i} - \left(X^{**i}\hat{\beta}^{**i}\right)^T\right)\left(Y^{**i} - X^{**i}\hat{\beta}^{**i}\right) \\ &= \left(Y^{**i^T} - \hat{\beta}^{**i^T} X^{**i^T}\right)\left(Y^{**i} - X^{**i}\hat{\beta}^{**i}\right) \\ e^{**i^T} e^{**i} &= X^{**i^T} Y^{**i} - Y^{**i^T} X^{**i}\hat{\beta}^{**i^T} - \hat{\beta}^{**i^T} Y^{**i^T} Y^{**i} + \hat{\beta}^{**i^T} X^{**i^T} Y \end{aligned} \tag{10}$$

The estimation results for the parameter $\hat{\beta}^{**i}$ are obtained by minimizing the sum of the squared errors, namely:

$$\begin{aligned} \frac{\partial\left(e^{**i^T} e^{**i}\right)}{\partial\hat{\beta}^{**i}} &= 0 \\ \frac{\partial\left(Y^{**i^T} Y^{**i}\right)}{\partial\hat{\beta}^{**i}} - 2\frac{\partial\left(Y^{**i^T} X^{**i}\hat{\beta}^{**i}\right)}{\partial\hat{\beta}^{**i}} + \frac{\partial\left(X^{**iT}\hat{\beta}^{**i} X^{**i}\hat{\beta}^{**i}\right)}{\partial\hat{\beta}^{**i}} &= 0 \\ -2Y^{**iT} X^{**i} + 2X^{**iT} X^{**i}\hat{\beta}^{**i} &= 0 \\ 2X^{**iT} X^{**i}\hat{\beta}^{**i} &= 2X^{**iT} X^{**i} \\ \hat{\beta}^{**i} &= \left(X^{**iT} X^{**i}\right)^{-1} Y^{**iT} X^{**i} \end{aligned} \tag{11}$$

So that the estimated value of $\hat{\beta}^{**i}$ is obtained as follows:

$$\hat{\beta}^{**i} = \left(X^{**iT} X^{**i}\right)^{-1} Y^{**iT} X^{**i}$$

The next step is to take the actual sample as in **Equation (9)**. Then the second row is omitted and the parameters are estimated using **Equation (11)**. Analogously applied to the third row to the $ke - n$. Then the Jackknife parameter $\hat{\beta}^{**1}, \hat{\beta}^{**2}, \cdots \hat{\beta}^{**n}$. Jackknife parameter estimation is obtained by finding the average value of each parameter estimator $\hat{\beta}^{1}, \hat{\beta}^{2}, \cdots \hat{\beta}^{n}$ as follows:

$$\hat{\beta}^{J} = \frac{\sum_{i=1}^{n} \hat{\beta}^{i}}{n} \tag{12}$$

Then calculate the level of accuracy of parameter estimation using bias and standard deviation. The bias of the Jackknife can be calculated as follows:

$$Bias^{J} = (n-1)\hat{\beta}^{J} - \hat{\beta} \tag{13}$$

The variance of the Jackknife can be calculated as follows:

$$Var(\hat{\beta}^{J}) = \frac{n-1}{n} \sum_{i=1}^{n} (\hat{\beta}^{i} - \hat{\beta}^{J})(\hat{\beta}^{i} - \hat{\beta}^{J})^{T} \tag{14}$$

So the Jackknife standard deviation is

$$SD^{J} = \sqrt{Var(\hat{\beta}^{J})} \tag{15}$$

2.5.2 Double Jackknife

The double Jackknife procedure is done by generating new data from the previously generated Jackknife data set. From the first stage Jackknife data set that was replicated as much as B_1 from the original data set, the Jackknifing process was repeated as many as B_2 replications, so that the total number of test statistics that had to be calculated as $B_1 + B_1 B_2$.

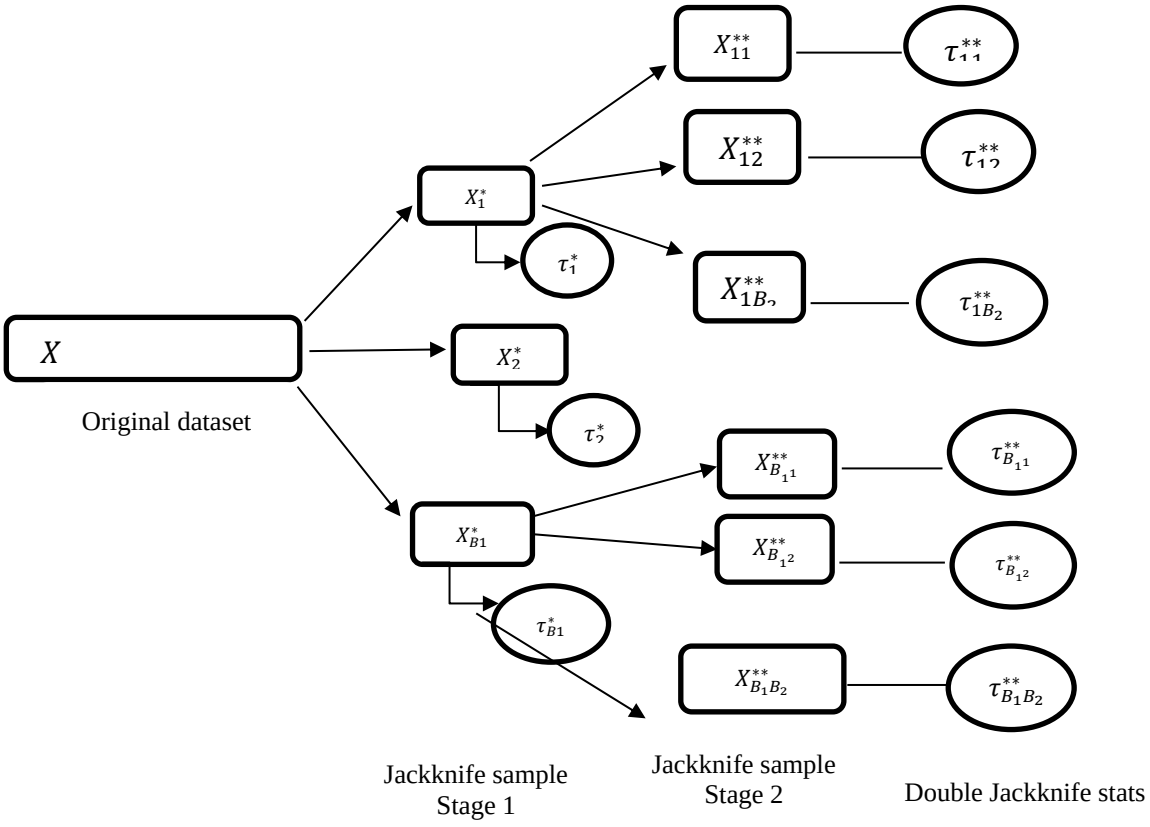

Figure 5. Double Jackknife procedure

The steps for the Double Jackknife resampling method are as follows:

1) Take a random sample of size n from the population and name it a vector $w_i = (Y_i, X_{ji})'$ with $Y_i = (y_1, y_2, \dots, y_n)'$ and $X_i = (x_1, x_2, \dots, x_n)'$ for $j = 1,2, \dots, k$ and $i = 1,2, \dots ., n$.
2) Delete the first line of the vector $w_i = (Y_i, X_{ji})'$ and it is a vector with sample size $(n-1)$ so vector y becomes $Y_i^{(j)} = \left(y_2^{(j)}, y_3^{(j)}, \dots, y_n^{(j)}\right)'$ and $X_{ji}^{(j)} = \left(x_{j2}^{(j)}, x_{j3}^{(j)}, \dots, x_{jn}^{(j)}\right)'$ as a sample with a Jackknife, remove one $\left(w_1^{(j)}\right)$ and estimate the path coefficient $\hat{\beta}^{(j1)}$ from $\left(w_1^{(j)}\right)$. Next, delete the second row of the vector $w_i = (Y_i, X_{ji})'$ and then label the sample size $(n-1)$ with the observation set $Y_i^{(j)} = \left(y_1^{(j)}, y_3^{(j)}, \dots, y_n^{(j)}\right)'$ and $X_{ji}^{(j)} = \left(x_{j1}^{(j)}, x_{j3}^{(j)}, \dots, x_{jn}^{(j)}\right)'$ as $\left(w_2^{(j)}\right)$ and estimate the path coefficient $\hat{\beta}^{(j2)}$. Remove one by one sample from n observations and estimate the path coefficient $\hat{\beta}^{(ji)}$ where the value $\hat{\beta}^{(ji)}$ is the Jackknife path coefficient. After eliminating the ith observation from $ke - i\ dari\ w_i$, the estimated value of the Jackknife path coefficient $\hat{\beta}^{(j1)}, \hat{\beta}^{(j2)}, \dots, \hat{\beta}^{(jn)}$.
3) Furthermore, the Double Jackknife procedure is carried out by regenerating data from the previously generated Jackknife data set. From the first phase of the Jackknife data set that was replicated as much as B_1 from the original data set, the Jackknife process was carried out again as many as B_2 replications, so that the total number of test statistics that had to be calculated as $B_1 + B_1 B_2$.
4) The Jackknife path coefficient is calculated which is the average of $\hat{\beta}^{(j1)}, \hat{\beta}^{(j2)}, \dots, \hat{\beta}^{(jn)}$.
5) Then calculate the level of accuracy of parameter estimation using bias and standard deviation.

The following are the steps that must be taken to estimate the standard error in a double jackknife:

1) Perform resampling by removing d observations alternately from the original sample set on each Jackknife sample
2) Perform parameter estimation on all Jackknife β samples
3) Calculates the standard error for the Jackknife sample

$$\hat{s}e_{jk} = \sqrt{\frac{n-d}{d\binom{n}{d}} \sum_{j=1}^{n} \left(\hat{\beta}_{(j)} - \frac{\hat{\beta}_j}{j}\right)^2} \tag{16}$$

2.6 Relative Efficiency

The comparison of the resampling method is measured based on the relative efficiency value **[20]**. 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.

$$eff\left(\hat{\beta}_{JK}, \hat{\beta}_{DJK}\right) = \frac{V(\hat{\beta}_{DJK})}{V(\hat{\beta}_{JK})} \tag{17}$$

Description:

$eff\left(\hat{\beta}_{JK}, \hat{\beta}_{DJK}\right)$ = Efficiency between Jackknife and Double jackknife resampling method estimator

$V(\hat{\beta}_{JK})$ = Variant of parameter estimator with the Jackknife resampling method

$V(\hat{\beta}_{DJK})$ = Variant of parameter estimator with Double Jackknife resampling method

Efficiency comparison between the Jackknife and Double Jackknife resampling estimator variants. That is a parameter estimator variant with the Double Jackknife resampling method. If the efficiency of the results using **Equation (17)** is more than 1, the $\hat{\beta}_{JK}$ an estimator is said to be more efficient than the $\hat{\beta}_{DJK}$ estimator. On the other hand, if the calculated efficiency results are less than 1, the $\hat{\beta}_{DJK}$ an estimator is said to be more efficient than the $\hat{\beta}_{JK}$. estimator. Meanwhile, if the efficiency results are equal to 1, both estimators are equally efficient.

3. RESULTS AND DISCUSSION

3.1 The Degree of Closeness of the Relationship between Categorical Variables Is Low

The level of closeness of the relationship between variables included in the low category is shown by the path coefficient values in the range of 0.05-0.20. Resampling was carried out on the results of parameter simulations in path analysis for each set of samples, the path coefficient was estimated, denoted as $\hat{\beta} = (\hat{\beta}_{XY_1} \quad \hat{\beta}_{XY_2} \quad \hat{\beta}_{Y_1Y_2})$. After obtaining three path coefficient estimators, namely $\hat{\beta}_{XY_1}$, $\hat{\beta}_{XY_2}$, and $\hat{\beta}_{Y_1Y_2}$ then the average path coefficient estimator is calculated, which is denoted as $\hat{\beta}^*_{XY_1}(.) \quad \hat{\beta}^*_{XY_2}(.)$, and $\hat{\beta}^*_{Y_1Y_2}(.)$. The hypothesis test is presented in **Table 1**.

Table 1. Hypothesis testing on conditions assuming normality not fulfilled and low closeness of Jackknife and Double Jackknife resampling

Variable	Jackknife Resampling		Double Jackknife Resampling		Relative Efficiency
	Path Coefficient	$p-value$	Path Coefficient	$p-value$	
	$n=100$	$n=100$	$n=100$	$n=100$	
$X \rightarrow Y_1$	0,01	0,00	0,01	0,00	1,04473
$X \rightarrow Y_2$	0,01	0,00	0,01	0,00	0,85679
$Y_1 \rightarrow Y_2$	0,92	0,00	0,93	0,00	0,83073

Based on **Table 1**, the p-value for the path coefficient, which shows the relationship between X to Y_1, X to Y_2 and Y_1 to Y_2 is smaller than the level set, so that H_0 is rejected. Thus, it can be concluded that X has a significant effect on Y_1, X has a significant effect on Y_2, and Y_1 has a significant effect on Y_2. Judging from the relative efficiency value of variable X to Y_1 in the path analysis, the efficiency value is more than one, indicating that the Jackknife method has a smaller variant than the Double Jackknife method. Whereas in the path analysis between variables X to Y_2 and Y_1 to Y_2 the efficiency value is less than one indicating that the Double Jackknife method has a smaller variant than the Jackknife method.

3.2 The Level of Closeness of the Relationship between the Variables in the Moderate Category

The level of closeness of the relationship between variables included in the medium category is shown by the path coefficient values in the range of 0.20-0.50. Resampling was carried out on the results of parameter simulations in path analysis for each set of samples, the path coefficient was estimated, denoted as $\hat{\beta} = (\hat{\beta}_{XY_1} \quad \hat{\beta}_{XY_2} \quad \hat{\beta}_{Y_1Y_2})$. After obtaining three path coefficient estimators, namely $\hat{\beta}_{XY_1}$, $\hat{\beta}_{XY_2}$ and $\hat{\beta}_{Y_1Y_2}$ then the average path coefficient estimator is calculated which is denoted as $\hat{\beta}^*_{XY_1}(.) \quad \hat{\beta}^*_{XY_2}(.)$, and $\hat{\beta}^*_{Y_1Y_2}(.)$. The hypothesis test is presented in **Table 2**.

Table 2. Hypothesis testing on conditions assuming normality not fulfilled and closeness of moderate Jackknife and Double Jackknife Resampling

Variable	Jackknife Resampling		Double Jackknife Resampling		Relative Efficiency
	Path Coefficient	$p-value$	Path Coefficient	$p-value$	
	$n=100$	$n=100$	$n=100$	$n=100$	
$X \rightarrow Y_1$	0,01	0,00	0,01	0,00	0,92524
$X \rightarrow Y_2$	0,00	0,00	0,00	0,00	0,96278
$Y_1 \rightarrow Y_2$	0,96	0,00	0,96	0,00	0,79389

Based on **Table 2**, the p-value for the path coefficient which shows the relationship between X to Y_1, X to Y_2 and Y_1 to Y_2 is smaller than the level set so that H_0 is rejected. Thus it can be concluded that X has a significant effect on Y_1, X has a significant effect on Y_2, and Y_1 has a significant effect on Y_2. Judging from

the relative efficiency values of all path analyses efficiency values are less than indicating that the Double Jackknife method has a smaller variant than the Jackknife method. Therefore, it can be concluded that path analysis with Double Jackknife resampling is more efficient than the Jackknife resampling method.

3.3 The Degree of Closeness of the Relationship between Category Variables Is High

The level of closeness of the relationship between variables included in the high category is shown by the path coefficient values in the range of 0.05-1.00. Resampling was carried out on the results of parameter simulations in path analysis for each set of samples, the path coefficient was estimated, denoted as $\hat{\beta} = (\hat{\beta}_{XY_1} \quad \hat{\beta}_{XY_2} \quad \hat{\beta}_{Y_1Y_2})$. After obtaining three path coefficient estimators, namely $\hat{\beta}_{XY_1}$, $\hat{\beta}_{XY_2}$ and $\hat{\beta}_{Y_1Y_2}$ then the average path coefficient estimator is calculated which is denoted as $\hat{\beta}^*_{XY_1}(.)$, $\hat{\beta}^*_{XY_2}(.)$, and $\hat{\beta}^*_{Y_1Y_2}(.)$. The hypothesis test is presented in **Table 3**.

Table 3. Hypothesis testing on unfulfilled normality assumption conditions and high relationship closeness of Resampling Jackknife and Double Jackknife

Variable	Jackknife Resampling		Double Jackknife Resampling		Relative Efficiency
	Path Coefficient	$p-value$	Path Coefficient	$p-value$	
	$n=100$	$n=100$	$n=100$	$n=100$	
$X \rightarrow Y_1$	0,01	0,00	0,01	0,00	0,99933
$X \rightarrow Y_2$	0,01	0,00	0,01	0,00	1,06331
$Y_1 \rightarrow Y_2$	0,95	0,00	0,93	0,00	1,40818

Based on **Table 3**, the p-value for the path coefficient, which shows the relationship between X to Y_1, X to Y_2, and Y_1 to Y_2, is smaller than the level set so that H_0 is rejected. Thus, it can be concluded that X has a significant effect on Y_1, X has a significant effect on Y_2, and Y_1 has a significant effect on Y_2. Judging from the relative efficiency value of variable X to Y_1, in path analysis, the efficiency value is less than one, indicating that the Double Jackknife method has a smaller variant than the Jackknife method, while in the path analysis between X to Y_2 and Y_1 to Y_1, the efficiency value is more than one, indicating that the method Jackknife has a smaller variant than the Double Jackknife method.

4 CONCLUSIONS

Based on the simulation studies that have been carried out, the use of the Jackknife and Double Jackknife resampling methods on the data with the assumption of residual normality is not met, indicating that both the Jackknife and Double Jackknife resampling methods can be applied and overcome normality. The calculated relative efficiency produces various levels of the closeness of the relationship between variables indicating that the Double Jackknife resampling method is more efficient than the Jackknife resampling method.

ACKNOWLEDGMENT

We thank Universitas Brawijaya for its support of this research.

REFERENCES

[1] J. C. Westland, "Partial least squares path analysis," *Struct. Equ. Model.*, pp. 17–38, 2019.
[2] R. J. Mitchell, "Path analysis: pollination," in *Design and analysis of ecological experiments*, Chapman and Hall/CRC, 2020, pp. 211–231.
[3] V. Y. Sundara, R. Warti, and A. Mardia, "SIMULASI METODE RESAMPLING DAN PENDUGAAN DATA HILANG TERBAIK," *J. Ris. dan Apl. Mat.*, vol. 3, no. 2, pp. 101–108, 2019.
[4] V. Chittora, "Resampling Techniques," *Biot. Res. Today*, vol. 4, no. 2, pp. 123–125, 2022.

[5] R. S. Dwiputra, "Perbandingan Resampling dengan Metode Jackknifing dan Bootstrapping pada Model WarpPLS." Universitas Brawijaya, 2018.
[6] A. Fernandes, "Comparison of Parameter Estimator Efficiency Levels of Path Analysis with Bootstrap and Jack Knife (Delete-5) Resampling Methods on Simulation Data," *J. Mat. Stat. dan Komputasi*, vol. 16, no. 3, pp. 353–364, 2020.
[7] J. Shao and D. Tu, *The jackknife and bootstrap*. Springer Science & Business Media, 2012.
[8] A. McIntosh, "The Jackknife estimation method," *arXiv Prepr. arXiv1606.00497*, 2016.
[9] A. Alifa, "Analisis Jalur Kuadratik dengan Resampling Bootstrap pada Data Simulasi." Universitas Brawijaya, 2020.
[10] Y. Y. Rinela, "Perbandingan Efisiensi Penduga dengan Pendekatan Resampling Bootstrap dan Blindfold pada Analisis Jalur (Studi pada Kualitas Produk Teh Casabat terhadap Kepuasan Pelanggan dan terhadap Loyalitas Pelanggan)." Universitas Brawijaya, 2018.
[11] U. N. Y. BINTI WASIO, "PERBANDINGAN EFEKTIVITAS METODE JACKKNIFE DAN METODE BOOTSTRAP DALAM ESTIMASI PARAMETER REGRESI DENGAN PENYIMPANGAN ASUMSI." Universitas Mataram, 2017.
[12] A. Fauzy, "Simulasi Tingkat Kepercayaan dari Data Berdistribusi Eksponensial Satu Parameter Tersensor Tipe-II Double," *Statistika*, vol. 14, no. 1, 2014.
[13] H. L. Nurjannah, "Konsistensi Besaran Resampling Jackknife pada Analisis Jalur Kuadratik." Universitas Brawijaya, 2020.
[14] W. van der Bijl, "phylopath: Easy phylogenetic path analysis in R," *PeerJ*, vol. 6, p. e4718, 2018.
[15] G. D. Garson, *Path analysis*. Statistical Associates Publishing Asheboro, NC, 2013.
[16] M. H. Afwa, "Efisiensi Penduga Parameter Teknik Resampling Bootstrap dan Jackknife pada Analisis Jalur Kuadratik (Studi Kasus Di Pemerintah Kota A)." Universitas Brawijaya, 2020.
[17] S. Dhaene and Y. Rosseel, "Resampling Based Bias Correction for Small Sample SEM," *Struct. Equ. Model. A Multidiscip. J.*, pp. 1–17, 2022.
[18] M. F. Tyas, "Studi Simulasi Efisiensi dan Konsistensi Resampling Bootstrap dan Jackknife dalam Menduga Parameter pada Analisis Jalur." Universitas Brawijaya, 2019.
[19] L. Wang and F. Yu, "Jackknife resampling parameter estimation method for weighted total least squares," *Commun. Stat. Methods*, vol. 49, no. 23, pp. 5810–5828, 2020.
[20] T. O. Prawita, "Perbandingan Tingkat Efisiensi Penduga Parameter Analisis Jalur dengan Resampling Bootstrap dan Jackknife Delete-5 pada Data Simulasi." Universitas Brawijaya, 2019.