Available online at http://scik.org
Commun. Math. Biol. Neurosci. 2022, 2022:35
https://doi.org/10.28919/cmbn/7258
ISSN: 2052-2541

PARAMETER ESTIMATION AND HYPOTHESIS TESTING THE SECOND ORDER OF BIVARIATE BINARY LOGISTIC REGRESSION (S-BBLR) MODEL WITH BERNDT HALL-HALL-HAUSMAN (BHHH) ITERATIONS

VITA RATNASARI[1,*], PURHADI[1], IGAR CALVERIA AVIANTHOLIB[1], ANDREA TRI RIAN DANI[2]

[1]Department of Statistics, Faculty of Science and Data Analytics, Institut Teknologi Sepuluh Nopember, Surabaya 60111, Indonesia

[2]Statistics Study Program, Department of Mathematics, Faculty of Mathematics and Natural Sciences, Universitas Mulawarman, Samarinda 75119, Indonesia

Abstract: Bivariate Binary Logistic Regression (BBLR) is a logistic model that has two response variables where each variable depends on two categories with the response variables being correlated with each other. In this research, a development study will be conducted on a Bivariate Binary Logistic Regression model using the second order (S-BBLR). Furthermore, the S-BBLR will be applied to the problem of Sustainable Development Goals (SDGs) related to the Human Development Index (HDI) and Public Health Development Index (PHDI) data in East Java, Indonesia. The parameter estimation process uses the Maximum Likelihood Estimator (MLE) method. The problem in estimate the parameters of this model is that MLE cannot find an implicit analytical solution, so an iteration method will be used in the form of Berndt Hall-Hall-Hausman (BHHH) in the iteration process. Hypothesis test for the S-BBLR model include simultaneous and partial tests performed using the Maximum Likelihood Ratio (MLRT) and the Wald method. Based on the analysis, it was found that the percentage of poor people, the pure participation rate (APM), and

*Corresponding author
E-mail address: vita_ratna@statistika.its.ac.id
Received February 10, 2022

the number of public health centers had a significant impact on PHI and PHDI with a classification accuracy of 86.84%.

Keywords: Berndt Hall-Hall-Hausman (BHHH); bivariate binary logistic regression (BBLR); human development index (HDI); maximum likelihood; poverty; public health development index (PHDI); sustainable development goals (SDGs); quality of life.

2010 AMS Subject Classification: 62P10.

1. INTRODUCTION

Analysis of regression is a statistical analysis method used to describe the relationship model between two or more variables [1]. In the relationship model, the variables used are grouped into two, namely response and predictor variables [2]. In general, regression analysis is grouped into four, namely Parametric Regression [3], Nonparametric Regression [4], Semiparametric Regression [5], and Logistic Regression [6]. In this study, logistic regression will be used.

Logistic regression is a regression model commonly used to model the relationship between a qualitative (category) response variable and one or more predictors [7], [8]. The logistic regression model can be used for the classification and prediction of response variables [9]. Modeling with logistic regression depends on the category and the number of categories on the response variable [10]. A logistic regression model that has a response variable with two categories is called a binary logistic regression model [9], [11]. Several previous studies have reviewed and developed binary logistic regression models, including: [7], [12]–[14].

The binary logistic regression model currently under development is limited to only one response variable every predictor variable used. In the application and in the real case, it is very likely that the binary logistic regression model will have multiple response variables. Furthermore, a binary logistic regression model using one response variable can be developed into a binary logistic regression model with two response variables called Bivariate Binary Logistics Regression (BBLR) [15]–[17]. There are several studies on the BBLR model that have been studied and developed by researchers. Ananth and Preisser [18] discussed the estimation of bivariate logistic regression parameters using the Maximum Likelihood Estimator (MLE). Bel, Fok, and Paap [19] discussed a multivariate logit model with three estimation methods, Stratified Importance Sampling, Composite Conditional Likelihood (CCL), and Generalized Method of Moments. And others have been completed by [15], [20]–[23].

The latest research that developed the First-order Bivariate Binary Logistic Regression (F-BBLR) model was carried out by [22]. A Bivariate binary logistic regression model with order one is one of a family of multivariate logit models that can be used to model the relationship between two binary response variables that are correlated with one or more predictor variables [15], [22], [24]. The bivariate binary logistic regression model, it has two binary responses that are correlated with the polynomial model, and the model response follows a multinomial distribution [22].

In this study, a Second-order Bivariate Binary Logistic Regression (S-BBLR) model will be developed. In the F-BBLR model, the predictor variable has degree one. Meanwhile, in the S-BBLR model, the predictor variable will be raised to the power of the quadratic or second-order. The underlying reason for this further study is to prove and explain that the second-order bivariate binary logistic regression (S-BBLR) model will produce better model accuracy in the classification process than the first-order bivariate binary logistic regression model. This is also supported by the real case in data that mostly has certain conditions, for example non-linear patterns.

The S-BBLR model applies to data on economic issues, namely the Human Development Index (HDI) and Public Health Development Index (PHDI) data in East Java, Indonesia. The human development index (HDI) is a comparative indicator of life expectancy, education, and living standards in all countries. The HDI is used as an indicator to assess aspects of the quality of development and to classify countries as developed, developing, and or underdeveloped countries [25]. The Ministry of Health creates an index, the Public Health Development Index (PHDI). PHDI is a collection of health indicators that can be easily and directly measured to explain health-related problems [26].

In this study, the parameter estimation method used is the Maximum Likelihood Estimator (MLE) with an iterative process using the Berndt Hall-Hall-Hausman (BHHH) method. Hypothesis testing to test the significance of the simultaneous parameters using the Maximum

Likelihood Ratio Test (MLRT) and to test the partial significance using the Wald method. The BHHH iteration method is a development and modification of the Fisher Scoring method, where the BHHH method will add the rule of many numbers. One of the advantages of the BHHH iteration method is that it only uses the first derivative [27]. It is hoped that the second-order bivariate binary logistic regression model proposed in this study, supported by the right variables, can be an alternative model capable of analyzing economic and health problems.

2. PRELIMINARIES

A. Contingency Table

A contingency table or what is often called cross-tabulation is a table that contains data on the number, frequency, or several classifications (category). In this study, a contingency table will be used. The contingency table is a classification of observation objects based on two variables and each variable is classified into two groups. Table 1 is a presentation of objects classified according to row variables (Factor A) and column variables (Factor B).

Factor A	***Factor B***		***Total***
	B_1	B_2	
A_1	n_{11}	n_{12}	$n_{1+} = n_{11} + n_{12}$
A_2	n_{21}	n_{22}	$n_{2+} = n_{21} + n_{22}$
Total	$n_{+1} = n_{11} + n_{21}$	$n_{+2} = n_{12} + n_{22}$	n

TABLE 1. Contingency Table of Factor A and Factor B

B. Multinomial Distribution

The multinomial distribution is a generalization of the binomial distribution into three or more categories. Suppose that each of the m predictor variables, identical studies can produce result in one of the c categories. If the result of trial i is category k, then $y_{jk} = 1$, otherwise $y_{jk} = 0$

otherwise, with $k = 1, 2, ..., c$ and $j = 1, 2, ..., m$. The multinomial distributions as follows:

$$P\left(Y_1 = y_1, Y_2 = y_2, ..., Y_{c-1} = y_{c-1}; m, \pi\right) = \begin{pmatrix} & m & \\ y_1 & y_2 & \dots & y_{c-1} \end{pmatrix} \pi_1^{y_1} \quad \pi_2^{y_2} \quad \dots \quad \pi_{c-1}^{y_{c-1}} \pi_c^{y_c} \tag{1}$$

where $\pi_c = 1 - \pi_1 - \pi_2 - ... - \pi_{c-1},\ y_c = m - y_1 - y_2 - ... - y_{c-1},\ y_j = 0, 1, 2, ..., m$.

$$E\left(Y_j\right) = m\pi_j, Var\left(Y_j\right) = m\pi_j\left(1 - \pi_j\right), Cov\left(Y_j, Y_{j^*}\right) \text{ for } j \neq j^* \tag{2}$$

For example, if $m = 1$. Equation (1) as follows:

$$P\left(Y_1 = y_1, Y_2 = y_2, ..., Y_{c-1} = y_{c-1}; 1, \pi\right) = \pi_1^{y_1} \pi_2^{y_2} ... \pi_{c-1}^{y_{c-1}} \pi_c^{y_c} \tag{3}$$

Where $y_1, y_2, ..., y_{c-1}$ is 0 or 1.

C. Independence Test

The Independence test is a test that is carried out to find out associations between categories by using χ^2 test statistics. Based on the two-dimensional contingency table, the hypothesis that underlies association testing are:

H_0 : no association between categories

H_1 : two categories are associated

with test statistics as follows:

$$Q_{count} = \sum_{i=1}^{2} \sum_{j=1}^{2} \frac{\left(n_{ij} - m_{ij}\right)^2}{m_{ij}}, \text{ where } m_{ij} = \frac{n_{i+} n_{+j}}{n_{++}} \tag{4}$$

Based on the test statistic, reject H_0 when $Q_{count} > \chi^2_{(1,\alpha)}$.

D. Binary Logistic Regression

Binary logistic regression is a data analysis method that is used to find relationships between the response variables that are binary, or whether response variables have two categories with a value of 0 or 1. If the response variable produces two categories, then the response variables follow the Bernoulli distribution, the probability function shows in Equation (5).

$$f(y)=\pi^{y}(1-\pi)^{1-y};\ y=0,1 \text{ then } f(y)=1-\pi \text{ and if } y=1 \text{ then } f(y)=\pi \quad (5)$$

The equation of the binary logistic regression model with k response variables can be expressed in Equation (6).

$$\pi(x)=\frac{\exp(\beta_0+\beta_1x_1+\beta_2x_2+...+\beta_kx_k)}{1+\exp(\beta_0+\beta_1x_1+\beta_2x_2+...+\beta_kx_k)} \quad (6)$$

The model in Equation (6) is transformed by a logit transformation, as follows:

$$g(x)=\ln\left(\frac{\pi(x)}{1-\pi(x)}\right)=\beta_0+\beta_1x_1+\beta_2x_2+...+\beta_kx_k \quad (7)$$

Given a random sample $Y_1,Y_2,...,Y_n$ with Bernoulli distribution with π parameter, then the model of logistic regression can be expressed as the following equation:

$$\text{logit}\left[\pi(x_i)\right]=\ln\left[\frac{\pi(x_i)}{1-\pi(x_i)}\right]=\boldsymbol{\beta}^T x_i \quad (8)$$

where:

$\boldsymbol{x}_i^T=\begin{bmatrix}1 & x_{1i} & x_{2i} & \cdots & x_{ki}\end{bmatrix}$

$\boldsymbol{\beta}^T=\begin{bmatrix}\beta_0 & \beta_1 & \beta_2 & \cdots & \beta_k\end{bmatrix}$

The estimation parameters use the Maximum Likelihood Estimation (MLE) method and the hypothesis test by the Maximum Likelihood Ratio Test (MLRT) method.

E. The First-Order Bivariate Binary Logistic Regression (F-BBLR)

The Bivariate Logistic Regression (BBLR) model is one of a family of multivariate logit models used to modeling the relationship between two binary responses that correlate with one or more predictor variables. Let Y_1 and Y_2 be two bivariate binary responses and $y=\begin{bmatrix}Y_{11} & Y_{10} & Y_{01}\end{bmatrix}^T$ is a vector of responses. The probabilities of observations for variables Y_1 and Y_2 are presented in Table 2.

	$Y_2 = 1$	$Y_2 = 0$	***Total***
$Y_1 = 1$	π_{11}	π_{10}	π_1
$Y_1 = 0$	π_{01}	π_{00}	$1-\pi_1$
Total	π_2	$1-\pi_2$	1

TABLE 2. Probabilities of Y_1 and Y_2

Based on [24], Table 2 are random variables $Y_{11}, Y_{10}, Y_{01},$ and Y_{00} . Because of $\pi_{11} + \pi_{10} + \pi_{01} + \pi_{00} = 1$ and $\pi_{00} = 1 - \pi_{11} - \pi_{10} - \pi_{01}$, then random variable $Y_{11}, Y_{10}, Y_{01},$ and Y_{00} are multinomial distribution with their probabilities $\pi_{11}, \pi_{10}, \pi_{01},$ and π_{00}.

Let $\boldsymbol{y} = [Y_{11} \quad Y_{10} \quad Y_{01}]^T$, so $\boldsymbol{y} \sim \boldsymbol{M}(1; \pi_{11}, \pi_{10}, \pi_{01})$. Based on Equation (8), the probability function $\boldsymbol{y}$ is defined as follows:

$$P(Y_{11} = y_{11}, Y_{10} = y_{10}, Y_{01} = y_{01}) = \prod_{g=0}^{1} \prod_{h=0}^{1} \pi_{gh}^{y_{gh}}, 0 < \pi_{gh} < 1 \tag{9}$$

$y_{gh} = 0,1; g, h = 0,1; y_{00} = 1 - y_{11} - y_{10} - y_{01};$ and $\pi_{00} = 1 - \pi_{11} - \pi_{10} - \pi_{01}$.

where:

y_{gh} is the value of Y_{gh} which represents the elements of the response vector

$\pi_{gh} = P(Y_1 = g, Y_2 = h)$ where are the marginal probabilities of Y_1 and Y_2 respectively

Let $\boldsymbol{x} = [1 \quad X_1 \quad X_2 \quad \dots \quad X_k]^T$ is a vector with dimensions $(k+1)$, so the bivariate binary logistic regression (BBLR) is given as follow:

$$\begin{aligned} \varphi_1(\boldsymbol{x}) &= \text{logit}(\pi_1(\boldsymbol{x})) \\ &= \ln\left(\frac{\pi_1(\boldsymbol{x})}{1-\pi_1(\boldsymbol{x})}\right) \\ &= \beta_{01} + \beta_{11}X_1 + \beta_{21}X_2 + \dots + \beta_{k1}X_k \end{aligned} \tag{10}$$

$$= \boldsymbol{x}^T \boldsymbol{\beta}_1$$

$$\varphi_2(\boldsymbol{x}) = \text{logit}(\pi_2(\boldsymbol{x}))$$

$$= \ln\left(\frac{\pi_2(\boldsymbol{x})}{1-\pi_2(\boldsymbol{x})}\right)$$

$$= \beta_{02} + \beta_{12}X_1 + \beta_{22}X_2 + \ldots + \beta_{k2}X_k$$

$$= \boldsymbol{x}^T \boldsymbol{\beta}_2 \tag{11}$$

$$\varphi_3(\boldsymbol{x}) = \ln(\psi(\boldsymbol{x}))$$

$$= \ln\left(\frac{\pi_{11}(\boldsymbol{x})\pi_{00}(\boldsymbol{x})}{\pi_{10}(\boldsymbol{x})\pi_{01}(\boldsymbol{x})}\right)$$

$$= \beta_{03} + \beta_{13}X_1 + \beta_{23}X_2 + \ldots + \beta_{k3}X_k$$

$$= \boldsymbol{x}^T \boldsymbol{\beta}_3 \tag{12}$$

Where $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \boldsymbol{\beta}_3$ is a parameter vector, $\pi_1(\boldsymbol{x})$ and $\pi_2(\boldsymbol{x})$ is a marginal probability of response variable, and $\psi(\boldsymbol{x})$ is the odds ratio of the response variable which shows that the response variables are correlated.

F. Berndt Hall-Hall-Hausman Algorithm

The Berndt Hall-Hall-Hausman (BHHH) method is a modification of the Fisher Scoring method, with modification I to H and for the information matrix is as follows:

$$I = -E\left[\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}^T \partial \boldsymbol{\theta}}\right] = \begin{bmatrix} -E\left(\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \theta_0^2}\right) & -E\left(\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \theta_0 \partial \theta_1}\right) & \cdots & -E\left(\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \theta_0 \partial \theta_k}\right) \\ -E\left(\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \theta_1 \partial \theta_0}\right) & -E\left(\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \theta_1^2}\right) & \cdots & -E\left(\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \theta_1 \partial \theta_k}\right) \\ \vdots & \vdots & \ddots & \vdots \\ -E\left(\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \theta_k \partial \theta_0}\right) & -E\left(\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \theta_k \partial \theta_1}\right) & \cdots & -E\left(\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \theta_k^2}\right) \end{bmatrix} \tag{13}$$

The Information matrix is modified to a Hessian matrix in Equation (14).

$$H(\boldsymbol{\beta}) = \left[\sum_{i=1}^{n} \left[\frac{\partial l_i(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}} \right]^T \frac{\partial l_i(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}^T} \right] \quad (14)$$

The iteration process using the B-HHH method is as follows:

$$\hat{\boldsymbol{\beta}}_{bhhh}^{(r+1)} = \hat{\boldsymbol{\beta}}_{bhhh}^{(r)} - H\left(\hat{\boldsymbol{\beta}}_{bhhh}^{(r)}\right)^{-1} g\left(\hat{\boldsymbol{\beta}}_{bhhh}^{(r)}\right), \text{ for } r = 0,1,2,... \quad (15)$$

Where $\hat{\boldsymbol{\beta}}_{bhhh}^{(r+1)}$ and $\hat{\boldsymbol{\beta}}_{bhhh}^{(r)}$ is the MLE parameter of the BLR model iteration $r+1$ and r. $H\left(\hat{\boldsymbol{\beta}}_{bhhh}^{(r)}\right)^{-1}$ show the inverse of the Hessian matrix of the MLE parameter of the BLR model at the $r-th$ iteration and $\boldsymbol{g}(\boldsymbol{\beta})$ is the first derivative of log-likelihood function. Iteration process will be stop if the convergence condition is $\left\| \hat{\boldsymbol{\beta}}^{(r+1)} - \hat{\boldsymbol{\beta}}^{(r)} \right\| \le \varepsilon$.

3. RESEARCH METHODOLOGY

A. Data

In this study the data that used are secondary data, published by Badan Pusat Statistik (BPS) in East Java Province and the Health Research and Development Agency (BALITBANGKES) of the Ministry of Health. The unit of observation used is all Regencies/Cities in East Java Province, up to 38 Regencies/Cities.

B. Research Variable

The variables used in this study are two response variables and three predictors, and presented in Table 3.

Variable	Research Variable	Variable Name	Category	Dummy
Response Variable	Y_1	Human Development Index (HDI)	***Moderate HDI*** = HDI values included in the interval 60 ≤ HDI < 70	0
			High HDI = HDI values included in the interval 70 ≤ HDI < 80	1

	Y_2	Public Health Development Index (PHDI)	***Low PHDI*** Low PHDI is a Regency/City that has an IPKM value less than the average IPKM in East Java	0
			High PHDI High PHDI is a Regency/City that has an IPKM more than or equal to the average IPKM in East Java.	1
Predictor Variable	X_1	Percentage of Poor Population	-	-
	X_2	The Pure Participation Rate for 13-15 Years Age Group	-	-
	X_3	The Number of Public Health Centers	-	-

TABLE 3. Research Variable

C. Analysis Procedure

Performing modeling for the second-order bivariate binary logistic regression (S-BBLR) model with the following stages. Suppose the first response variable is denoted by Y_1 and the second response variable is Y_2, then the S-BBLR model can be expressed as follows:

$$\begin{aligned}
\eta_1(\boldsymbol{x}_i^*) &= \text{logit}\left(\pi_1^*(\boldsymbol{x}_i^*)\right) \\
&= \ln\left(\frac{\pi_1^*(\boldsymbol{x}_i^*)}{1-\pi_1^*(\boldsymbol{x}_i^*)}\right) \\
&= \beta_{10} + \beta_{11}x_{1i} + \beta_{12}x_{2i} + \ldots + \beta_{1k}x_{ki} + \beta_{111}x_{1i}^2 + \beta_{122}x_{2i}^2 + \ldots + \beta_{1kk}x_{ki}^2 + \beta_{112}x_{1i}x_{2i} + \\
&\quad \beta_{113}x_{1i}x_{3i} + \ldots + \beta_{11k}x_{1i}x_{ki} + \beta_{123}x_{2i}x_{3i} + \beta_{124}x_{2i}x_{4i} + \ldots + \beta_{12k}x_{2i}x_{ki} + \ldots + \beta_{1,k-1,k}x_{ki-1}x_{ki}
\end{aligned}$$

$$\eta_1(\boldsymbol{x}_i^*) = \boldsymbol{x}_i^{*T}\boldsymbol{\beta}_1$$

$$\begin{aligned}
\eta_2(\boldsymbol{x}_i^*) &= \text{logit}\left(\pi_2^*(\boldsymbol{x}_i^*)\right) \\
&= \ln\left(\frac{\pi_2^*(\boldsymbol{x}_i^*)}{1-\pi_2^*(\boldsymbol{x}_i^*)}\right) \\
&= \beta_{20} + \beta_{21}x_{1i} + \beta_{22}x_{2i} + \ldots + \beta_{2k}x_{ki} + \beta_{211}x_{1i}^2 + \beta_{222}x_{2i}^2 + \ldots + \beta_{2kk}x_{ki}^2 + \beta_{212}x_{1i}x_{2i} \\
&\quad + \beta_{213}x_{1i}x_{3i} + \ldots + \beta_{21k}x_{1i}x_{ki} + \beta_{223}x_{2i}x_{3i} + \beta_{224}x_{2i}x_{4i} + \ldots + \beta_{22k}x_{2i}x_{ki} + \ldots + \beta_{2,k-1,k}x_{ki-1}x_{ki}
\end{aligned}$$

$$\eta_2(\boldsymbol{x}) = \boldsymbol{x}_i^{*T}\boldsymbol{\beta}_2$$

$$
\begin{aligned}
\eta_3(\boldsymbol{x}_i^*) &= \ln\left(\psi(\boldsymbol{x}_i^*)\right) \\
&= \ln\left(\frac{\pi_{11}^*\left(\boldsymbol{x}_i^*\right)\pi_{00}^*\left(\boldsymbol{x}_i^*\right)}{\pi_{10}^*\left(\boldsymbol{x}_i^*\right)\pi_{01}^*\left(\boldsymbol{x}_i^*\right)}\right) \\
&= \beta_{30} + \beta_{31}x_{1i} + \beta_{32}x_{2i} + \ldots + \beta_{3k}x_{ki} + \beta_{311}x_{1i}^2 + \beta_{322}x_{2i}^2 + \ldots + \beta_{3kk}x_{ki}^2 + \beta_{312}x_{1i}x_{2i} \\
&\quad + \beta_{313}x_{1i}x_{3i} + \ldots + \beta_{31k}x_{1i}x_{ki} + \beta_{323}x_{2i}x_{3i} + \beta_{324}x_{2i}x_{4i} + \ldots + \beta_{32k}x_{2i}x_{ki} + \ldots + \beta_{3,k-1,k}x_{ki-1}x_{ki} \\
\eta_3(\boldsymbol{x}_i^*) &= \boldsymbol{x}_i^{*T}\boldsymbol{\beta}_3
\end{aligned}
$$

where:

$\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \boldsymbol{\beta}_3$ is a vector parameter.

$$\boldsymbol{\beta}_1 = \left[\beta_{10} \quad \beta_{11} \quad \beta_{12} \quad \cdots \quad \beta_{1k} \quad \beta_{111} \quad \beta_{122} \quad \cdots \quad \beta_{1kk} \quad \beta_{112} \quad \beta_{113} \quad \cdots \quad \beta_{11k} \quad \beta_{123} \quad \beta_{124} \quad \cdots \quad \beta_{12k} \quad \beta_{1,k-1,k}\right]^T$$

$$\boldsymbol{\beta}_2 = \left[\beta_{20} \quad \beta_{21} \quad \beta_{22} \quad \cdots \quad \beta_{2k} \quad \beta_{211} \quad \beta_{222} \quad \cdots \quad \beta_{2kk} \quad \beta_{212} \quad \beta_{213} \quad \cdots \quad \beta_{21k} \quad \beta_{223} \quad \beta_{224} \quad \cdots \quad \beta_{22k} \quad \beta_{2,k-1,k}\right]^T$$

$$\boldsymbol{\beta}_3 = \left[\beta_{30} \quad \beta_{31} \quad \beta_{32} \quad \cdots \quad \beta_{3k} \quad \beta_{311} \quad \beta_{322} \quad \cdots \quad \beta_{3kk} \quad \beta_{312} \quad \beta_{313} \quad \cdots \quad \beta_{31k} \quad \beta_{323} \quad \beta_{324} \quad \cdots \quad \beta_{32k} \quad \beta_{3,k-1,k}\right]^T$$

$\pi_1^*\left(\boldsymbol{x}_i^*\right)$ and $\pi_2^*\left(\boldsymbol{x}_i^*\right)$ is a marginal probability of response variable, and $\psi\left(\boldsymbol{x}_i^*\right)$ is the odds ratio of the response variables which shows that the response variables are correlated.

$$P\left(Y_{1i} = 1 \middle| \boldsymbol{x}_i^*\right) = \pi_1^*(\boldsymbol{x}_i^*) = \frac{\exp\left(x_i^{*T}\boldsymbol{\beta}_1\right)}{1+\exp\left(x_i^{*T}\boldsymbol{\beta}_1\right)}$$

$$P\left(Y_{1i} = 0 \middle| \boldsymbol{x}_i^*\right) = 1 - \pi_1^*(\boldsymbol{x}_i^*) = \frac{1}{1+\exp\left(x_i^{*T}\boldsymbol{\beta}_1\right)}$$

$$P\left(Y_{2i} = 1 \middle| \boldsymbol{x}_i^*\right) = \pi_2^*(\boldsymbol{x}_i^*) = \frac{\exp\left(x_i^{*T}\boldsymbol{\beta}_2\right)}{1+\exp\left(x_i^{*T}\boldsymbol{\beta}_2\right)}$$

$$P\left(Y_{2i} = 0 \middle| \boldsymbol{x}_i^*\right) = 1 - \pi_2^*(\boldsymbol{x}_i^*) = \frac{1}{1+\exp\left(x_i^{*T}\boldsymbol{\beta}_2\right)}$$

The joint probability density function of the variables Y_{1i} and Y_{2i} is:

$$P\left(Y_{11i} = y_{11i}, Y_{10i} = y_{10i}, Y_{01i} = y_{01i}\right) = \left(\pi_{11}^*\left(\boldsymbol{x}_i^*\right)\right)^{y_{11i}} \left(\pi_{10}^*\left(\boldsymbol{x}_i^*\right)\right)^{y_{10i}} \left(\pi_{01}^*\left(\boldsymbol{x}_i^*\right)\right)^{y_{01i}} \left(\pi_{00}^*\left(\boldsymbol{x}_i^*\right)\right)^{y_{00i}} \qquad (16)$$

1. Calculating parameter estimation using the MLE Method

a. The forming likelihood function with the MLE method.

$$L(\boldsymbol{\beta}|\boldsymbol{y}) = \prod_{i=1}^{n} f(\boldsymbol{y}_i|\boldsymbol{\beta}) \tag{17}$$

b. Maximize the likelihood function

$$l(\boldsymbol{\beta}) = \ln L(\boldsymbol{\beta}) \tag{18}$$

c. Determine the first partial derivative of the log-likelihood function to the parameter.

$$g(\boldsymbol{\beta}) = \left[\left[\frac{\partial l(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}_1^T} \right]^T \quad \left[\frac{\partial l(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}_2^T} \right]^T \quad \left[\frac{\partial l(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}_3^T} \right]^T \right]^T \tag{19}$$

d. If a non-closed-form equation is obtained, then we use BHHH iterations with the formula:

$$\hat{\boldsymbol{\beta}}_{bhhh}^{(r+1)} = \hat{\boldsymbol{\beta}}_{bhhh}^{(r)} - \boldsymbol{H}\left(\hat{\boldsymbol{\beta}}_{bhhh}^{(r)}\right)^{-1} \boldsymbol{g}\left(\hat{\boldsymbol{\beta}}_{bhhh}^{(r)}\right), \text{ for } r = 0,1,2,... \tag{20}$$

with:

$$g(\boldsymbol{\beta}) = \left[\left[\frac{\partial l(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}_1^T} \right]^T \quad \left[\frac{\partial l(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}_2^T} \right]^T \quad \left[\frac{\partial l(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}_3^T} \right]^T \right]^T$$

$$\boldsymbol{H}(\boldsymbol{\beta}) = \left[\sum_{i=1}^{n} \frac{\partial^2 l_i(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}\, \partial \boldsymbol{\beta}^T} \right]$$

$$\frac{\partial^2 l_i(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}\, \partial \boldsymbol{\beta}^T} = \begin{bmatrix} \frac{\partial^2 l_i(\boldsymbol{\beta})}{\partial \beta_0^2} & \frac{\partial^2 l_i(\boldsymbol{\beta})}{\partial \beta_0 \partial \beta_1} & \cdots & \frac{\partial^2 l_i(\boldsymbol{\beta})}{\partial \beta_0 \partial \beta_k} \\ \frac{\partial^2 l_i(\boldsymbol{\beta})}{\partial \beta_1 \partial \beta_0} & \frac{\partial^2 l_i(\boldsymbol{\beta})}{\partial \beta_1^2} & \cdots & \frac{\partial^2 l_i(\boldsymbol{\beta})}{\partial \beta_1 \partial \beta_k} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 l_i(\boldsymbol{\beta})}{\partial \beta_k \partial \beta_0} & \frac{\partial^2 l_i(\boldsymbol{\beta})}{\partial \beta_k \partial \beta_1} & \cdots & \frac{\partial^2 l_i(\boldsymbol{\beta})}{\partial \beta_k^2} \end{bmatrix},$$

2. Simultaneous hypothesis testing

a. Determine the test hypothesis.

$$H_0 : \beta_{h1} = \beta_{h2} = ... = \beta_{hk} = \beta_{h11} = \beta_{h22} = ... = \beta_{hkk} = \beta_{h12} = \beta_{h13} = ... =$$

$$\beta_{h1k} = \beta_{h23} = \beta_{h24} = ... = \beta_{h24} = ... = \beta_{h2k} = \beta_{h,k-1,k} = 0 \quad h = 1,2,3$$

H_1 : there is at least one $\beta_{gh} \neq 0$, $g = 1,2,...,k$ $h = 1,2,3$

b. Determine the parameter set and the likelihood function under the population and under H_0.

c. Forming and maximizing the log-likelihood function under the population and under H_0.

d. Forming Odds Ratio.

$$\Lambda = \frac{L(\hat{\omega})}{L(\hat{\Omega})} \tag{21}$$

e. Determine $G^2 = -2\ln\Lambda = 2\left[\ln L(\hat{\Omega}) - \ln L(\hat{\omega})\right]$.

f. Determine the distribution G^2 and rejection area H_0.

3. Partial hypothesis testing

a. Determine the test hypothesis

H_0 : $\beta_{gh} = 0$

H_1 : $\beta_{gh} \neq 0$, $g = 1,2,...,k$ $h = 1,2,3$

b. Determine Z test statistics and rejection area H_0.

4. MAIN RESULTS

4.1 Theoretical Study

In this section, it is explained the results of research regarding parameter estimation and hypothesis testing.

A. The Second-Order Bivariate Binary Logistic Regression (S-BBLR) Model Parameter Estimation

The second-order bivariate binary logistic regression (S-BBLR) model has $3(k+1)$ parameters, with $(k+1)$ regression parameter which shows a correlation between the response variables and $2(k+1)$ regression parameter which shows a relationship between predictor variables and response variables. The estimation of the parameters of the S-BBLR model that cannot be obtained analytically will be approached with the iteration process using BHHH.

Let $\boldsymbol{y}_i = \begin{bmatrix} Y_{1i} & Y_{2i} \end{bmatrix}^T = \begin{bmatrix} Y_{11i} & Y_{10i} & Y_{01i} \end{bmatrix}^T, i = 1, 2, ..., n$ be samples of the same independent and identical random vectors with a multinomial distribution shown by $\boldsymbol{y}_i \sim M\left(1, \pi_{11}^*\left(\boldsymbol{x}_i^*\right), \pi_{10}^*\left(\boldsymbol{x}_i^*\right), \pi_{01}^*\left(\boldsymbol{x}_i^*\right), \pi_{00}^*\left(\boldsymbol{x}_i^*\right)\right)$ where $\pi_{11}^*\left(\boldsymbol{x}_i^*\right), \pi_{10}^*\left(\boldsymbol{x}_i^*\right), \pi_{01}^*\left(\boldsymbol{x}_i^*\right), \pi_{00}^*\left(\boldsymbol{x}_i^*\right)$ is the probability of each random variable $Y_{11i}, Y_{10i}, Y_{01i}$, and Y_{00i} that load $\boldsymbol{\beta}$. The joint probability distribution fuction defined as:

$$\begin{aligned} f\left(\boldsymbol{y}_i \mid \boldsymbol{\beta}\right) &= P\left(Y_{11i} = y_{11i}, Y_{10i} = y_{10i}, Y_{01i} = y_{01i}, Y_{00i} = y_{00i}\right) \\ &= \left(\pi_{11}^*\left(\boldsymbol{x}_i^*\right)\right)^{y_{11i}} \left(\pi_{10}^*\left(\boldsymbol{x}_i^*\right)\right)^{y_{10i}} \left(\pi_{01}^*\left(\boldsymbol{x}_i^*\right)\right)^{y_{01i}} \left(\pi_{00}^*\left(\boldsymbol{x}_i^*\right)\right)^{y_{00i}} \end{aligned} \tag{22}$$

Based on the probability function in Equation (22), then the form the likelihood function as follows:

$$\begin{aligned} L_i\left(\boldsymbol{\beta}\right) &= \prod_{i=1}^{n} f\left(\boldsymbol{y}_i \mid \boldsymbol{\beta}\right) \\ &= \prod_{i=1}^{n} P\left(Y_{11i} = y_{11i}, Y_{10i} = y_{10i}, Y_{01i} = y_{01i}\right) \\ &= \left(\pi_{11}^*\left(\boldsymbol{x}_i^*\right)\right)^{y_{11i}} \left(\pi_{10}^*\left(\boldsymbol{x}_i^*\right)\right)^{y_{10i}} \left(\pi_{01}^*\left(\boldsymbol{x}_i^*\right)\right)^{y_{01i}} \left(\pi_{00}^*\left(\boldsymbol{x}_i^*\right)\right)^{y_{00i}} \end{aligned} \tag{23}$$

Let $\left(\pi_{gh}^*\left(\boldsymbol{x}_i^*\right)\right)^{y_{gh}} = \left(\pi_{ghi}^*\right)^{y_{ghi}}$ for $g, h = 0,1$; $i = 1, 2, ..., n$, so the likelihood function in Equation (24).

$$L_i\left(\boldsymbol{\beta}\right) = \prod_{i=1}^{n} \left(\pi_{11}^*\right)^{y_{11i}} \left(\pi_{10}^*\right)^{y_{10i}} \left(\pi_{01}^*\right)^{y_{01i}} \left(\pi_{00}^*\right)^{y_{00i}} \tag{24}$$

Futhermore, the likelihood function is easier to maximize in the form $\ln L_i\left(\boldsymbol{\beta}\right)$ and expressed by $l_i\left(\boldsymbol{\beta}\right)$.

$$\begin{aligned} l_i\left(\boldsymbol{\beta}\right) &= \ln L_i\left(\boldsymbol{\beta}\right) \\ &= \sum_{i=1}^{n} \left(y_{11i} \ln \pi_{11i}^* + y_{10i} \ln \pi_{10i}^* + y_{01i} \ln \pi_{01i}^* + y_{00i} \ln \pi_{00i}^*\right) \end{aligned} \tag{25}$$

Based on the definition [28], the gradient vector of the log-likelihood function in Equation (25) as follows:

$$\boldsymbol{g}(\boldsymbol{\beta}) = \left[\left[\frac{\partial l(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}_1^T} \right]^T \quad \left[\frac{\partial l(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}_2^T} \right]^T \quad \left[\frac{\partial l(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}_3^T} \right]^T \right]^T \tag{26}$$

Let $\eta_1 = \eta_1(\boldsymbol{x}_i^*), \eta_2 = \eta_2(\boldsymbol{x}_i^*)$, and $\eta_3 = \eta_3(\boldsymbol{x}_i^*)$, so it takes the form of $\boldsymbol{\eta} = [\eta_1 \quad \eta_2 \quad \eta_3]$; $\boldsymbol{\pi} = [\pi_{11}^* \quad \pi_{10}^* \quad \pi_{01}^* \quad \pi_{00}^*]^T$. Then, determine the vector derivative from $\boldsymbol{\eta}$ to $\boldsymbol{\pi}^*$, that is $\dfrac{\partial \boldsymbol{\eta}}{\partial \boldsymbol{\pi}^*}$.

Because of the vector $\boldsymbol{\pi}^*$ has four elements, while vectors $\boldsymbol{\eta}$ only have three elements, to get $\dfrac{\partial \boldsymbol{\eta}}{\partial \boldsymbol{\pi}^*}$ symmetrical, for example $\eta_0 = \ln \pi_{++}^*$, with $\pi_{++}^* = \pi_{11}^* + \pi_{10}^* + \pi_{01}^* + \pi_{00}^*$. So that it is obtained $\eta = [\eta_0 \quad \eta_1 \quad \eta_2 \quad \eta_3]^T$. Let D_1 is a matrix whose elements are vectors $\boldsymbol{\eta}$ to $\boldsymbol{\pi}^*$:

$$\boldsymbol{D}_1 = \begin{bmatrix} \frac{\partial \eta_0}{\partial \pi_{11}^*} & \frac{\partial \eta_0}{\partial \pi_{10}^*} & \frac{\partial \eta_0}{\partial \pi_{01}^*} & \frac{\partial \eta_0}{\partial \pi_{00}^*} \\ \frac{\partial \eta_1}{\partial \pi_{11i}^*} & \frac{\partial \eta_1}{\partial \pi_{10i}^*} & \frac{\partial \eta_1}{\partial \pi_{01i}^*} & \frac{\partial \eta_1}{\partial \pi_{00i}^*} \\ \frac{\partial \eta_2}{\partial \pi_{11i}^*} & \frac{\partial \eta_2}{\partial \pi_{10i}^*} & \frac{\partial \eta_2}{\partial \pi_{01i}^*} & \frac{\partial \eta_2}{\partial \pi_{00i}^*} \\ \frac{\partial \eta_3}{\partial \pi_{11i}^*} & \frac{\partial \eta_3}{\partial \pi_{10i}^*} & \frac{\partial \eta_3}{\partial \pi_{01i}^*} & \frac{\partial \eta_3}{\partial \pi_{00i}^*} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ \frac{1}{\pi_{1i}^*} & \frac{1}{\pi_{1i}^*} & -\frac{1}{1-\pi_{1i}^*} & -\frac{1}{1-\pi_{1i}^*} \\ \pi_{2i}^* & -\frac{1}{1-\pi_{2i}^*} & \frac{1}{\pi_{2i}^*} & -\frac{1}{1-\pi_{2i}^*} \\ \frac{1}{\pi_{11i}^*} & -\frac{1}{\pi_{10i}^*} & -\frac{1}{\pi_{01i}^*} & \frac{1}{\pi_{00i}^*} \end{bmatrix}. \tag{27}$$

Based on Equation (27) obtained the inverse of the matrix as follows:

$$\boldsymbol{D}_1^{-1} = \begin{bmatrix} \pi_{11i}^* & \frac{\pi_{11i}^* \pi_{01i}^*}{\pi_{1i}^* \Delta_{1i}} & \frac{\pi_{11i}^* \pi_{10i}^*}{\pi_{1i}^* \Delta_{1i}} & \Delta_i \\ \pi_{10i}^* & \frac{\pi_{10i}^* \pi_{00i}^*}{(1-\pi_{2i}^*)\Delta_{1i}} & -\frac{\pi_{11i}^* \pi_{10i}^*}{\pi_{1i}^* \Delta_{1i}} & -\Delta_i \\ \pi_{01i}^* & -\frac{\pi_{11i}^* \pi_{01i}^*}{\pi_{2i}^* \Delta_{1i}} & \frac{\pi_{01i}^* \pi_{00i}^*}{(1-\pi_{1i}^*)\Delta_{1i}} & -\Delta_i \\ \pi_{00i}^* & -\frac{\pi_{10i}^* \pi_{00i}^*}{(1-\pi_{2i}^*)\Delta_{1i}} & -\frac{\pi_{01i}^* \pi_{00i}^*}{(1-\pi_{1i}^*)\Delta_{1i}} & \Delta_i \end{bmatrix}, \tag{28}$$

where $\Delta_{1i} = \dfrac{\pi_{11i}^* \pi_{10i}^* \pi_{01i}^* \pi_{00i}^*}{\pi_{1i}^* (1-\pi_{1i}^*) \pi_{2i}^* (1-\pi_{2i}^*) \Delta_i}$ and $\Delta_i = \left(\dfrac{1}{\pi_{11i}^*} + \dfrac{1}{\pi_{10i}^*} + \dfrac{1}{\pi_{01i}^*} + \dfrac{1}{\pi_{00i}^*} \right)^{-1}$.

Gradient vector from the log-likelihood function in the Equation (25) can be write as:

$$\boldsymbol{g}_i(\boldsymbol{\beta}) = \frac{\partial l_i(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}}. \tag{29}$$

Based on equations (24) to (27), the gradient vector elements are obtained in the following equation:

$$\boldsymbol{g}_i(\boldsymbol{\beta}) = \frac{\partial l_i(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}_1} = \frac{\partial\left(y_{11i}\ln \pi^*_{11i} + y_{10i}\ln \pi^*_{10i} + y_{01i}\ln \pi^*_{01i} + y_{00i}\ln \pi^*_{00i}\right)}{\partial \boldsymbol{\beta}_1}$$

$$= \frac{y_{11i}}{\pi^*_{11i}}\frac{\partial \pi^*_{11i}}{\partial \boldsymbol{\beta}_1} + \frac{y_{10i}}{\pi^*_{10i}}\frac{\partial \pi^*_{10i}}{\partial \boldsymbol{\beta}_1} + \frac{y_{01i}}{\pi^*_{01i}}\frac{\partial \pi^*_{01i}}{\partial \boldsymbol{\beta}_1} + \frac{y_{00i}}{\pi^*_{00i}}\frac{\partial \pi^*_{00i}}{\partial \boldsymbol{\beta}_1}$$

$$\boldsymbol{g}_i(\boldsymbol{\beta}) = \frac{1}{\Delta_{1i}}\left(\frac{y_{11i}\pi^*_{01i} - y_{01i}\pi^*_{11i}}{\pi^*_{2i}} + \frac{y_{10i}\pi^*_{00i} - y_{00i}\pi^*_{10i}}{1-\pi^*_{2i}}\right)\boldsymbol{x}_i \tag{30}$$

$$\boldsymbol{g}_i(\boldsymbol{\beta}) = \frac{\partial l_i(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}_2} = \frac{\partial\left(y_{11i}\ln \pi^*_{11i} + y_{10i}\ln \pi^*_{10i} + y_{01i}\ln \pi^*_{01i} + y_{00i}\ln \pi^*_{00i}\right)}{\partial \boldsymbol{\beta}_2}$$

$$= \frac{y_{11i}}{\pi^*_{11i}}\frac{\partial \pi^*_{11i}}{\partial \boldsymbol{\beta}_2} + \frac{y_{10i}}{\pi^*_{10i}}\frac{\partial \pi^*_{10i}}{\partial \boldsymbol{\beta}_2} + \frac{y_{01i}}{\pi^*_{01i}}\frac{\partial \pi^*_{01i}}{\partial \boldsymbol{\beta}_2} + \frac{y_{00i}}{\pi^*_{00i}}\frac{\partial \pi^*_{00i}}{\partial \boldsymbol{\beta}_2}$$

$$\boldsymbol{g}_i(\boldsymbol{\beta}) = \frac{1}{\Delta_{1i}}\left(\frac{y_{11i}\pi^*_{10i} - y_{10i}\pi^*_{11i}}{\pi^*_{1i}} + \frac{y_{01i}\pi^*_{00i} - y_{00i}\pi^*_{01i}}{1-\pi^*_{1i}}\right)\boldsymbol{x} \tag{31}$$

$$\boldsymbol{g}_i(\boldsymbol{\beta}) = \frac{\partial l_i(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}_3} = \frac{\partial\left(y_{11i}\ln \pi^*_{11i} + y_{10i}\ln \pi^*_{10i} + y_{01i}\ln \pi^*_{01i} + y_{00i}\ln \pi^*_{00i}\right)}{\partial \boldsymbol{\beta}_3}$$

$$= \frac{y_{11i}}{\pi^*_{11i}}\frac{\partial \pi^*_{11i}}{\partial \boldsymbol{\beta}_3} + \frac{y_{10i}}{\pi^*_{10i}}\frac{\partial \pi^*_{10i}}{\partial \boldsymbol{\beta}_3} + \frac{y_{01i}}{\pi^*_{01i}}\frac{\partial \pi^*_{01i}}{\partial \boldsymbol{\beta}_3} + \frac{y_{00i}}{\pi^*_{00i}}\frac{\partial \pi^*_{00i}}{\partial \boldsymbol{\beta}_3}$$

$$\boldsymbol{g}_i(\boldsymbol{\beta}) = \Delta_{1i}\left(\frac{y_{11i}}{\pi^*_{11i}} - \frac{y_{10i}}{\pi^*_{10i}} - \frac{y_{01i}}{\pi^*_{01i}} + \frac{y_{10i}}{\pi^*_{00i}}\right)\boldsymbol{x}_i \tag{32}$$

With $\Delta_{1i} = \frac{\pi^*_{11i}\pi^*_{10i}\pi^*_{01i}\pi^*_{00i}}{\pi^*_{1i}\left(1-\pi^*_{1i}\right)\pi^*_{2i}\left(1-\pi^*_{2i}\right)\Delta_i}$ and $\Delta_i = \left(\frac{1}{\pi^*_{11i}} + \frac{1}{\pi^*_{10i}} + \frac{1}{\pi^*_{01i}} + \frac{1}{\pi^*_{00i}}\right)^{-1}$, $i = 1, 2, \ldots, n$.

Based on [28], there is a relationship between the gradient vector and the Hessian matrix:

$$E\left(\boldsymbol{g}(\boldsymbol{\beta})\right) = 0 \; ; \; Var\left(\boldsymbol{g}(\boldsymbol{\beta})\right) = E\left(\boldsymbol{g}(\boldsymbol{\beta})\boldsymbol{g}^T(\boldsymbol{\beta})\right). \tag{33}$$

Furthermore, the Hessian matrix has a relationship with the Information matrix:

$$\boldsymbol{I}(\boldsymbol{\beta}) = -\boldsymbol{H}(\boldsymbol{\beta}) \tag{34}$$

Meanwhile [29] show that there is a relationship between the gradient vector and the information matrix:

$$Var\left(\boldsymbol{g}\left(\boldsymbol{\beta}\right)\right)=n\boldsymbol{I}\left(\boldsymbol{\beta}\right) \tag{35}$$

Based on Equation (33) to (35), the Hessian matrix is obtained as follows:

$$\begin{aligned}
\boldsymbol{H}\left(\boldsymbol{\beta}\right) &= \left[\sum_{i=1}^{n}\left[\frac{\partial l_i\left(\boldsymbol{\beta}\right)}{\partial \boldsymbol{\beta}^T}\right]^T \frac{\partial l_i\left(\boldsymbol{\beta}\right)}{\partial \boldsymbol{\beta}^T}\right] \\
&= \left\{\left[\sum_{i=1}^{n}\frac{1}{\Delta_{1i}}\left(\frac{y_{11i}\pi_{01i}^{*}-y_{01i}\pi_{11i}^{*}}{\pi_{2i}^{*}}+\frac{y_{10i}\pi_{00i}^{*}-y_{00i}\pi_{10i}^{*}}{1-\pi_{2i}^{*}}\right)\boldsymbol{x}_i+\frac{1}{\Delta_{1i}}\right.\right. \\
&\left.\left(\frac{y_{11i}\pi_{10i}^{*}-y_{10i}\pi_{11i}^{*}}{\pi_{1i}^{*}}+\frac{y_{01i}\pi_{00i}^{*}-y_{00i}\pi_{01i}^{*}}{1-\pi_{1i}^{*}}\right)\boldsymbol{x}_i+\Delta_{1i}\left(\frac{y_{11i}}{\pi_{11i}^{*}}-\frac{y_{10i}}{\pi_{10i}^{*}}-\frac{y_{01i}}{\pi_{01i}^{*}}+\frac{y_{10i}}{\pi_{00i}^{*}}\right)\boldsymbol{x}_i\right] \\
&\left[\frac{1}{\Delta_{1i}}\left(\frac{y_{11i}\pi_{01i}^{*}-y_{01i}\pi_{11i}^{*}}{\pi_{2i}^{*}}+\frac{y_{10i}\pi_{00i}^{*}-y_{00i}\pi_{10i}^{*}}{1-\pi_{2i}^{*}}\right)\boldsymbol{x}_i+\frac{1}{\Delta_{1i}}\right. \\
&\left.\left.\left(\frac{y_{11i}\pi_{10i}^{*}-y_{10i}\pi_{11i}^{*}}{\pi_{1i}^{*}}+\frac{y_{01i}\pi_{00i}^{*}-y_{00i}\pi_{01i}^{*}}{1-\pi_{1i}^{*}}\right)\boldsymbol{x}_i+\Delta_{1i}\left(\frac{y_{11i}}{\pi_{11i}^{*}}-\frac{y_{10i}}{\pi_{10i}^{*}}-\frac{y_{01i}}{\pi_{01i}^{*}}+\frac{y_{10i}}{\pi_{00i}^{*}}\right)\boldsymbol{x}_i\right]\right\}
\end{aligned} \tag{36}$$

After obtaining the gradient vector and the Hessian matrix, it is possible to perform a numerical iteration process using the Berndt Hall-Hall-Hausman (B-HHH) method to obtained the MLE of the S-BBLR model parameters.

$$\hat{\boldsymbol{\beta}}_{bhhh}^{(r+1)}=\hat{\boldsymbol{\beta}}_{bhhh}^{(r)}-\boldsymbol{H}\left(\hat{\boldsymbol{\beta}}_{bhhh}^{(r)}\right)^{-1}\boldsymbol{g}\left(\hat{\boldsymbol{\beta}}_{bhhh}^{(r)}\right),\ \text{for } r=0,1,2,\ldots \tag{37}$$

Where $\hat{\boldsymbol{\beta}}_{bhhh}^{(r)}$ and $\hat{\boldsymbol{\beta}}_{bhhh}^{(r+1)}$ is the MLE parameter of the S-BBLR model at the r-th and $r+1$ iterations, $\boldsymbol{g}\left(\hat{\boldsymbol{\beta}}_{bhhh}^{(r)}\right)$ is the gradient vector for the S-BBLR model parameter MLE at the r-th iteration, and $\boldsymbol{H}\left(\hat{\boldsymbol{\beta}}_{bhhh}^{(r)}\right)^{-1}$ is the inverse of the Hessian matrix. The iteration process will stop if the convergence condition is $\left\|\hat{\boldsymbol{\beta}}_{bhhh}^{(r+1)}-\hat{\boldsymbol{\beta}}_{bhhh}^{(r)}\right\|\leq\varepsilon$, where ε is a very small positive number.

B. Hypothesis Testing

Hypothesis testing on the S-BBLR model consists of a simultaneous testing and a partial testing. Testing the parameters of the S-BBLR model simultaneously using the Maximum

Likelihood Ratio Test (MLRT) method. The hypotheses used for the simultaneous test are:

H_0 : $\beta_{h1}=\beta_{h2}=...=\beta_{hk}=\beta_{h11}=\beta_{h22}=...=\beta_{hkk}=\beta_{h12}=\beta_{h13}=...=$

$\beta_{h1k}=\beta_{h23}=\beta_{h24}=...=\beta_{h24}=...=\beta_{h2k}=\beta_{h,k-1,k}=0 \;\; h=1,2,3$

H_1 : there is at least one $\beta_{gh}\neq 0$, $g=1,2,...,k$ $h=1,2,3$

Determination of test statistics using the MLRT method begins with determining the set of model parameters below H_0: $\omega_2=\{\beta_{01},\beta_{02},\beta_{03}\}$. Next, form the likelihood function below H_0:

$$L(\omega_2)=\prod_{i=1}^{n}\left(\left(\pi_{11i}^{*}\right)^{y_{11i}}\left(\pi_{10i}^{*}\right)^{y_{10i}}\left(\pi_{01i}^{*}\right)^{y_{01i}}\left(\pi_{00i}^{*}\right)^{y_{00i}}\right) \tag{38}$$

Let $\hat{\boldsymbol{\beta}}^{*}=\begin{bmatrix}\hat{\beta}_{01} & \hat{\beta}_{02} & \hat{\beta}_{03}\end{bmatrix}$ is the MLE for the parameter H_0, then the likelihood function is obtained below H_0:

$$\begin{aligned}L\left(\hat{\boldsymbol{\beta}}^{*}\right)&=\max_{\boldsymbol{\beta}^{*}\in\omega_2}L(\omega_2)\\&=\prod_{i=1}^{n}\left(\left(\pi_{11i}^{*}\right)^{y_{11i}}\left(\pi_{10i}^{*}\right)^{y_{10i}}\left(\pi_{01i}^{*}\right)^{y_{01i}}\left(\pi_{00i}^{*}\right)^{y_{00i}}\right)\end{aligned} \tag{39}$$

The log-likelihood function below H_0, as follows:

$$\begin{aligned}l(\omega_2)&=\ln L(\omega_2)\\&=\sum_{i=1}^{n}\left(y_{11i}\ln\pi_{11i}^{*}+y_{10i}\ln\pi_{10i}^{*}+y_{01i}\ln\pi_{01i}^{*}+y_{00i}\ln\pi_{00i}^{*}\right)\end{aligned} \tag{40}$$

Determine the maximum log-likelihood function by:

$$\frac{\partial l(\hat{\omega}_2)}{\partial\beta_{01}}=\sum_{i=1}^{n}\left\{\frac{y_{11i}}{\hat{\pi}_{11i}}\frac{\hat{\pi}_{11i}}{\partial\beta_{01}}+\frac{y_{10i}}{\hat{\pi}_{10i}}\frac{\hat{\pi}_{10i}}{\partial\beta_{01}}+\frac{y_{01i}}{\hat{\pi}_{01i}}\frac{\hat{\pi}_{01i}}{\partial\beta_{01}}+\frac{y_{00i}}{\hat{\pi}_{00i}}\frac{\hat{\pi}_{00i}}{\partial\beta_{01}}\right\}$$

$$0=\sum_{i=1}^{n}\left\{\left(\frac{y_{11i}}{\hat{\pi}_{11i}}\frac{\hat{\pi}_{11i}\hat{\pi}_{01i}}{\hat{\pi}_{2i}\Delta_i}1\right)+\left(\frac{y_{10i}}{\hat{\pi}_{10i}}\frac{\hat{\pi}_{10i}\hat{\pi}_{00i}}{(1-\hat{\pi}_{2i})\Delta_i}1\right)+\left(\frac{y_{01i}}{\hat{\pi}_{01i}}\left(-\frac{\hat{\pi}_{11i}\hat{\pi}_{01i}}{\hat{\pi}_{2i}\Delta_i}1\right)\right)+\right\}$$

$$=\left\{\left(\frac{y_{00i}}{\hat{\pi}_{00i}}\left(-\frac{\hat{\pi}_{10i}\hat{\pi}_{00i}}{(1-\hat{\pi}_{2i})\Delta_i}1\right)\right)\right\}$$

$$\frac{\partial l(\hat{\omega}_2)}{\partial \beta_{02}} = \sum_{i=1}^{n} \left\{ \frac{y_{11i}}{\hat{\pi}_{11i}} \frac{\hat{\pi}_{11i}}{\partial \beta_{02}} + \frac{y_{10i}}{\hat{\pi}_{10i}} \frac{\hat{\pi}_{10i}}{\partial \beta_{02}} + \frac{y_{01i}}{\hat{\pi}_{01i}} \frac{\hat{\pi}_{01i}}{\partial \beta_{02}} + \frac{y_{00i}}{\hat{\pi}_{00i}} \frac{\hat{\pi}_{00i}}{\partial \beta_{02}} \right\}$$

$$0 = \sum_{i=1}^{n} \left\{ \left(\frac{y_{11i}}{\hat{\pi}_{11i}} \frac{\hat{\pi}_{11i}\hat{\pi}_{01i}}{\hat{\pi}_{2i}\Delta_i} 1 \right) + \left(\frac{y_{10i}}{\hat{\pi}_{10i}} \frac{\hat{\pi}_{10i}\hat{\pi}_{00i}}{(1-\hat{\pi}_{2i})\Delta_i} 1 \right) + \left(\frac{y_{01i}}{\hat{\pi}_{01i}} \left(-\frac{\hat{\pi}_{11i}\hat{\pi}_{01i}}{\hat{\pi}_{2i}\Delta_i} 1 \right) \right) + \right\}$$

$$= \left\{ \left(\frac{y_{00i}}{\hat{\pi}_{00i}} \left(-\frac{\hat{\pi}_{10i}\hat{\pi}_{00i}}{(1-\hat{\pi}_{2i})\Delta_i} 1 \right) \right) \right\}$$

$$\frac{\partial l(\hat{\omega}_2)}{\partial \beta_{03}} = \sum_{i=1}^{n} \left\{ \frac{y_{11i}}{\hat{\pi}_{11i}} \frac{\hat{\pi}_{11i}}{\partial \beta_{03}} + \frac{y_{10i}}{\hat{\pi}_{10i}} \frac{\hat{\pi}_{10i}}{\partial \beta_{03}} + \frac{y_{01i}}{\hat{\pi}_{01i}} \frac{\hat{\pi}_{01i}}{\partial \beta_{03}} + \frac{y_{00i}}{\hat{\pi}_{00i}} \frac{\hat{\pi}_{00i}}{\partial \beta_{03}} \right\}$$

$$0 = \sum_{i=1}^{n} \left\{ \left(\frac{y_{11i}}{\hat{\pi}_{11i}} - \frac{y_{10i}}{\hat{\pi}_{10i}} - \frac{y_{01i}}{\hat{\pi}_{01i}} + \frac{y_{00i}}{\hat{\pi}_{00i}} \right) \left(\frac{1}{\hat{\pi}_{11i}} + \frac{1}{\hat{\pi}_{10i}} + \frac{1}{\hat{\pi}_{01i}} + \frac{1}{\hat{\pi}_{00i}} \right)^{-1} \right\}$$

where:

$$\Delta_{1i} = \frac{\pi^*_{11i}\pi^*_{10i}\pi^*_{01i}\pi^*_{00i}}{\pi^*_{1i}(1-\pi^*_{1i})\pi^*_{2i}(1-\pi^*_{2i})\Delta_i} \text{ and } \Delta_i = \left(\frac{1}{\pi^*_{11i}} + \frac{1}{\pi^*_{10i}} + \frac{1}{\pi^*_{01i}} + \frac{1}{\pi^*_{00i}} \right)^{-1}, \; i = 1, 2, \ldots, n.$$

$$\hat{\pi}^*_{11i} = \begin{cases} \dfrac{\left(a_{2i} - \sqrt{a_{2i}^2 + b_{2i}}\right)}{2(\psi_i - 1)}, \psi_i \neq 1 \\ \hat{\pi}^*_{1i}\hat{\pi}^*_{2i} \quad, \psi_i = 1 \end{cases}$$

with $a_{2i} = 1 + (\hat{\pi}^*_1 + \hat{\pi}^*_2)(\psi_i - 1)$, $b = -4\psi_i(\psi_i - 1)\hat{\pi}^*_1\hat{\pi}^*_2$.

$$\psi_i = \frac{\hat{\pi}^*_{11i}\hat{\pi}^*_{00i}}{\hat{\pi}^*_{10i}\hat{\pi}^*_{01i}}, \quad \hat{\pi}^*_{1i} = \frac{\exp(\hat{\beta}_{01})}{1 + \exp(\hat{\beta}_{01})}, \text{ and } \hat{\pi}^*_{2i} = \frac{\exp(\hat{\beta}_{02})}{1 + \exp(\hat{\beta}_{02})},$$

$$\hat{\pi}^*_{10i} = \hat{\pi}^*_{1i} - \hat{\pi}^*_{11i}$$
$$\hat{\pi}^*_{01i} = \hat{\pi}^*_{2i} - \hat{\pi}^*_{11i}$$
$$\hat{\pi}^*_{00i} = 1 - \hat{\pi}^*_{1i} - \hat{\pi}^*_{2i} + \hat{\pi}^*_{11i}$$

After getting the maximum likelihood function below H_0, then determining the set of model parameters under the population, namely:

$$\Omega_2 = \{\beta_{1k}, \beta_{1kk}, \beta_{11k}, \beta_{12k} ..., \beta_{1,k-1,k}, \beta_{2k}, \beta_{2kk}, \beta_{21k}, \beta_{22k} ..., \beta_{2,k-1,k}, \beta_{3k}, \beta_{3kk}, \beta_{31k}, \beta_{32k} ..., \beta_{3,k-1,k}\}.$$

Next, form the likelihood function under the population:

$$L(\Omega_2) = \prod_{i=1}^{n} \left(\pi_{11i}^{y_{11i}} \pi_{10i}^{y_{10i}} \pi_{01i}^{y_{01i}} \pi_{00i}^{y_{00i}} \right) \tag{41}$$

For example, $\hat{\boldsymbol{\beta}} = \begin{bmatrix} \hat{\beta}_1^T & \hat{\beta}_2^T & \hat{\beta}_3^T \end{bmatrix}^T$ is the MLE for the parameter in the population, then the maximum likelihood function under the population:

$$\begin{aligned} L(\hat{\boldsymbol{\beta}}) &= \max_{\boldsymbol{\beta} \in \Omega_2} L(\Omega_2) \\ &= \prod_{i=1}^{n} \left(\hat{\pi}_{11i}^{y_{11i}} \hat{\pi}_{10i}^{y_{10i}} \hat{\pi}_{01i}^{y_{01i}} \hat{\pi}_{00i}^{y_{00i}} \right) \end{aligned} \tag{42}$$

where:

$$\hat{\pi}_{11i}^* = \begin{cases} \dfrac{\left(a_{3i} - \sqrt{a_{3i}^2 + b_{3i}}\right)}{2(\psi_i - 1)}, \psi_i \neq 1 \\ \hat{\pi}_{1i}^* \hat{\pi}_{2i}^* \quad , \psi_i = 1 \end{cases}$$

with $a_{3i} = 1 + (\hat{\pi}_{1i} + \hat{\pi}_{2i})(\psi_i - 1)$, $b_{3i} = -4\psi_i (\psi_i - 1) \hat{\pi}_{1i} \hat{\pi}_{2i}$

$$\psi_i = \frac{\hat{\pi}_{11i} \hat{\pi}_{00i}}{\hat{\pi}_{10i} \hat{\pi}_{01i}}, \quad \hat{\pi}_{1i}^* = \frac{\exp\left(\hat{\beta}_1^T x_i^*\right)}{1 + \exp\left(\hat{\beta}_1^T x_i^*\right)}, \text{ dan } \hat{\pi}_{2i}^* = \frac{\exp\left(\hat{\beta}_2^T x_i^*\right)}{1 + \exp\left(\hat{\beta}_2^T x_i^*\right)},$$

$$\begin{aligned} \hat{\pi}_{10i} &= \hat{\pi}_{1i} - \hat{\pi}_{11i} \\ \hat{\pi}_{01i} &= \hat{\pi}_{2i} - \hat{\pi}_{11i} \\ \hat{\pi}_{00i} &= 1 - \hat{\pi}_{1i} - \hat{\pi}_{2i} + \hat{\pi}_{11i} \end{aligned}$$

Based on Equations (39) to (42), the odds ratio is obtained as follows:

$$\Lambda_2 = \frac{L(\hat{\boldsymbol{\beta}}^*)}{L(\hat{\boldsymbol{\beta}})} = \frac{\prod_{i=1}^{n} \left(\left(\pi_{11i}^*\right)^{y_{11i}} \left(\pi_{10i}^*\right)^{y_{10i}} \left(\pi_{01i}^*\right)^{y_{01i}} \left(\pi_{00i}^*\right)^{y_{00i}} \right)}{\prod_{i=1}^{n} \left(\hat{\pi}_{11i}^{y_{11i}} \hat{\pi}_{10i}^{y_{10i}} \hat{\pi}_{01i}^{y_{01i}} \hat{\pi}_{00i}^{y_{00i}} \right)}. \tag{43}$$

Next, determine the text statistics for the simultaneous test hypothesis in Equation (44):

$$G_2^2 = -\ln\Lambda_2 = -\ln\left(\frac{L\left(\hat{\boldsymbol{\beta}}^*\right)}{L\left(\hat{\boldsymbol{\beta}}\right)}\right)^2 = 2\left[\ln L\left(\hat{\boldsymbol{\beta}}\right) - \ln L\left(\hat{\boldsymbol{\beta}}^*\right)\right] \tag{44}$$

For the large samples, G_2^2 statistics can be approximated by a Chi-Square distribution with df degree of freedom, where df is the difference between the number of parameters under the population and the number of parameters under H_0. Rejected area of H_0 is $G_2^2 > \chi^2_{(\alpha,df)}$.

If the simultaneous test result is rejected H_0, run a partial test to determine the effect of the predictor variables on the response variable individually. The hypothesis in the partial test are as follows:

H_0 : $\beta_{gh} = 0$

H_1 : $\beta_{gh} \neq 0$, $g = 1,2,...,k$ $h = 1,2,3$

The test statistics for hypothesis testing in the partial test is the Wald test in Equation (45).

$$Z = \frac{\hat{\theta}_{gh}}{\sqrt{\hat{\mathrm{V}}\mathrm{ar}\left(\hat{\theta}_{gh}\right)}} \underset{n\to\infty}{\sim} N(0,1) \tag{45}$$

Where $\hat{\mathrm{V}}\mathrm{ar}\left(\hat{\theta}_{gh}\right)$ obtained from the diagonal element to $(gh+1)$ from the covariance-variance matrix $\mathrm{cov}\left(\hat{\boldsymbol{\theta}}\right)$, $\hat{\mathrm{c}}\mathrm{ov}\left(\hat{\boldsymbol{\theta}}\right) = \left[I\left(\hat{\boldsymbol{\theta}}\right)\right]^{-1} = -\left[H\left(\hat{\boldsymbol{\theta}}\right)\right]^{-1}$. The rejected area of H_0 is $|Z| > Z_{\alpha/2}$.

4.2 Data Application

A. Description of Research Data

The variables in this study consist of two response variables, HDI (Y_1) with $Y_1 = 0$ for the moderate HDI and $Y_1 = 1$ for the high HDI. The PHDI (Y_2) with $Y_2 = 0$ for the low PHDI and $Y_2 = 1$ for the high PHDI.

	$Y_2 = 1$	$Y_2 = 0$	***Total***
$Y_1 = 1$	19	1	20
$Y_1 = 0$	8	10	18
Total	27	11	38

TABLE 4. Contingency Table of HDI and PHDI

B. Correlation

Before conducting an analysis using Bivariate Binary Logistic Regression (BBLR) model, it is important to first examine the correlation value between the response variables and also the correlation value between the predictor variables.

1. Correlation between Response Variables

To measure the correlation between response variables, the odds ratio value is used (ψ). The criterion used is if the odds ratio (ψ) between the response variables is $\psi > 1$, so it can be concluded if there is a positive correlation between the response variables.

Y_1	Y_2		ψ	CI 95% for ψ
	$Y_2 = 1$	$Y_2 = 0$		
$Y_1 = 1$	19	1	23.7500	$2.5914 \le \psi \le 217.6691$
$Y_1 = 0$	8	10		

TABLE 5. Correlation Value between Response Variables

Based on Table 5, it can be concluded that there is a positive correlation between the variables HDI (Y_1) and PHDI (Y_2).

2. *Correlation between Predictor Variables*

Furthermore, multicollinearity testing between predictor variables will be carried out using the Variance Inflation Factor (VIF) value.

Predictor Variables	**VIF**
X_1	1.2034
X_2	1.1574
X_3	1.0439

TABLE 6. Multicollinearity Test between Predictor Variables

Based on Table 6, it can be seen that all predictor variables ($X_1 - X_3$) have a VIF value of less than 10, so it can be concluded that there is no multicollinearity problem between predictor variables.

C. *The First-Order Bivariate Binary Logistic Regression (F-BBLR)*

In this section, an analysis will be carried out using the First-Order Bivariate Binary Logistics Regression (F-BBLR) model. The F-BBLR model is a non-spatial model (global model), so each location is assumed to be homogeneous.

1. *Parameter Estimation*

The estimation of the parameters in the F-BBLR model using the B-HHH iteration is shown in Table 7.

Variable	**Parameter**	**Estimate**	**Standard Error**	**Z-value**
Intercept	β_{10}	-4.9740	9.2150e-08	-5.3977e+07
X_1	β_{11}	-0.6271	9.3434e-07	-6.7114e+05
X_2	β_{12}	0.1657	8.1077e-06	2.0437e+04

Variable	Parameter	Estimate	Standard Error	Z-value
X_3	β_{13}	-0.2169	4.5947e-06	-4.7197e+04
Intercept	β_{20}	7.6812	1.3379e-08	5.7411e+08
X_1	β_{21}	-0.3027	1.5179e-07	-1.9943e+06
X_2	β_{22}	-0.0052	1.1863e-06	-4.3660e+03
X_3	β_{23}	0.1430	6.1852e-07	2.3119e+05
Intercept	β_{30}	-4.9736	1.8548e-11	-2.6814e+11
X_1	β_{31}	-0.5816	1.9739e-10	-2.9463e+09
X_2	β_{32}	0.1685	1.6692e-09	1.0093e+08
X_3	β_{33}	-0.0147	7.1771e-10	-2.0481e+07

TABLE 7. The Results of Parameter Estimation by F-BBLR model

Based on the parameter estimation results in Table 7, we can calculate the classification accuracy of the F-BBLR model. The classification accuracy in this study used the Apparent Error Rate (APER) criteria. This APER value indicates the proportion of observations that are misclassified by the F-BBLR model.

$$1-APER=1-0.2631=0.7369 \tag{46}$$

The accuracy of the classification of the F-BBLR model in classifying HDI (Y_1) and PHDI (Y_2) in the Regency/City of East Java Province is 73.69%.

2. *Hypothesis Testing*

In the F-BBLR model, parameters are tested simultaneously and partially. The hypothesis used for the simultaneous test is as follows:

H_0 : $\beta_{h1} = \beta_{h2} = \beta_{h3} = 0$; $h = 1,2,3$

H_1 : there is at least one $\beta_{gh} \neq 0$; $g = 1,2,3$ $h = 1,2,3$

The result of the calculation, the value of test statistics G^2 is 205.0470. From the *chi-square* distribution table obtained the value of $\chi^2_{(0,1;9)}$ is 16.6480. Because of the value of G^2 is more than $\chi^2_{(0,1;9)}$, so the decision is rejected H_0. It can be concluded that at least, there is one predictor variables which significant effect on HDI and PHDI in the Regency/City of East Java Province.

Furthermore, partial parameter testing is carried out to find out significant predictor variables on the model with the tested hypothesis are:

H_0 : $\beta_{gh} = 0$

H_1 : $\beta_{gh} \neq 0$, $g = 1,2,3$ $h = 1,2,3$

Based on Table 7, it can be seen that the value of the test statistics $|Z|$ for all estimates parameters more than $Z_{(0.1/2)}$, so the decision is rejected H_0 and it can be concluded that all predictor variables have a significant effect on the F-BBLR model on level significance $\alpha = 10\%$.

D. The Second-Order Bivariate Binary Logistic Regression (S-BBLR)

1. Parameter Estimation

The estimation of the parameters in the S-BBLR model using the B-HHH iteration is shown in Table 8.

Variable	**Parameter**	**Estimate**	**Standard Error**	**Z-value**
Intercept	β_{10}	-320.8553	6.2103e-19	-5.1665e+20
X_1	β_{11}	0.4276	9.9406e-18	4.3014e+16
X_2	β_{12}	7.4275	6.0121e-17	1.2354e+17
X_3	β_{13}	-2.4389	1.9149e-17	-1.2736+17

Variable	Parameter	Estimate	Standard Error	Z-value
X_1X_1	β_{111}	0.0075	1.5426e-16	4.8582e+13
X_2X_2	β_{122}	-0.042	5.8342e-15	-6.8375e+12
X_3X_3	β_{133}	0.0042	5.5319e-16	7.6026e+12
X_1X_2	β_{112}	-0.0162	9.6476e-16	-1.6756e+13
X_1X_3	β_{113}	0.0137	2.7808e-16	4.9201e+13
X_2X_3	β_{123}	0.0226	1.8443e-15	1.2247e+13
Intercept	β_{20}	-337.6334	2.6280e-15	-1.2842e+17
X_1	β_{21}	12.1478	2.9693e-14	4.0912e+14
X_2	β_{22}	7.4284	2.4801e-13	2.9952e+13
X_3	β_{23}	-6.5615	6.5230e-14	-1.0059e+14
X_1X_1	β_{211}	-0.0739	3.4613e-13	-2.1358e+11
X_2X_2	β_{222}	-0.0403	2.34093e-11	-1.7218e+09
X_3X_3	β_{233}	1.2550e-04	1.6447e-12	7.6302e+07
X_1X_2	β_{212}	-0.1382	2.7939e-12	-4.9462e+10
X_1X_3	β_{213}	0.0763	7.3547e-13	1.0374e+11
X_2X_3	β_{223}	0.0650	6.1472e-12	1.0568e+10
Intercept	β_{30}	-347.2447	1.5628e-22	-2.2219e+24
X_1	β_{31}	10.8319	2.4776e-21	4.3718e+21
X_2	β_{32}	7.2744	1.3459e-20	5.4048e+20

Variable	Parameter	Estimate	Standard Error	Z-value
X_3	β_{33}	-4.5231	5.1737e-21	-8.7425e+20
X_1X_1	β_{311}	-0.0481	3.9552e-20	-1.2173e+18
X_2X_2	β_{322}	-0.0361	1.1601e-18	-3.1135e+16
X_3X_3	β_{333}	0.0018	1.5225e-19	1.1664e+16
X_1X_2	β_{312}	-0.1276	2.1010e-19	-6.0730e+17
X_1X_3	β_{313}	0.0640	8.4207e-20	7.5989e+17
X_2X_3	β_{323}	0.0424	4.4966e-19	9.4383e+16

TABLE 8. The Results of Parameter Estimation by S-BBLR

Based on the parameter estimation results from the S-BBLR model, we can calculate the classification accuracy. The same thing as in the F-BBLR model, in this study using the Apparent Error Rate (APER) criteria. The APER value was obtained from the S-BBLR model.

$$1 - APER = 1 - 0.1315 = 0.8684 \qquad (47)$$

The classification accuracy of the S-BBLR model in classifying HDI and PHDI in the Regency/City of East Java Province is 86.84%.

2. *Hypothesis Testing*

The results of simultaneous hypothesis testing on the S-BBLR model were obtained, with the G^2 value statistic of 276.9620. From the *chi-square* distribution obtained the value of $\chi^2_{(0,1;27)}$ is 36.7410. Because of the value of G^2 is more than $\chi^2_{(0,1;27)}$, so the decision is reject H_0.

Furthermore, partial parameter testing is carried out to find out significant predictor variables on the model S-BBLR. Based on Table 8, it can be seen that the value of the test statistics $|Z|$ for all estimates parameters more than $Z_{(0.1/2)}$, so the decision is reject H_0 and it can be concluded that all predictor variables have a significant effect on the S-BBLR model on level significance $\alpha = 10\%$.

E. Selection of the Best Model

Based on the results in Table 7 and 8, it can be seen that the S-BBLR model with the B-HHH iteration method produces better classification accuracy than the F-BBLR model with the same iteration method. The S-BBLR model can be written for modeling HDI and PHDI in the Regency/City of East Java Province as follows:

$$\hat{\eta}_1(\boldsymbol{x}) = \ln\left(\frac{\hat{\pi}_1(\boldsymbol{x})}{1-\hat{\pi}_1(\boldsymbol{x})}\right)$$
$$= -320.8553 + 0.4276X_1 + 7.4275X_2 - 2.4389X_3 + 0.0075X_1^2 -$$
$$0.042X_2^2 + 0.0042X_3^2 - 0.0162X_1X_2 + 0.0137X_1X_3 + 0.0226X_2X_3$$

$$\hat{\eta}_2(\boldsymbol{x}) = \ln\left(\frac{\hat{\pi}_2(\boldsymbol{x})}{1-\hat{\pi}_2(\boldsymbol{x})}\right)$$
$$= -337.6334 + 12.1478X_1 + 7.4284X_2 - 6.5615X_3 + 0.0075X_1^2 -$$
$$0.042X_2^2 + 0.0042X_3^2 - 0.0162X_1X_2 + 0.0763X_1X_3 + 0.065X_2X_3$$

$$\eta_3(\boldsymbol{x}) = \ln\left(\frac{\hat{\pi}_{11}(\boldsymbol{x})\hat{\pi}_{00}(\boldsymbol{x})}{\hat{\pi}_{10}(\boldsymbol{x})\hat{\pi}_{01}(\boldsymbol{x})}\right)$$
$$= -347.2447 + 10.8319X_1 + 7.2744X_2 + 0.0018X_3 - 0.048X_1^2 -$$
$$0.0361X_2^2 + 0.0018X_3^2 - 0.1276X_1X_2 + 0.064X_1X_3 + 0.0424X_2X_3$$

In the S-BBLR model, the accuracy of the classification is 86.84%.

5. CONCLUSION

The study of parameter estimation and hypothesis testing on the second-order bivariate binary logistic regression (S-BBLR) model which was then applied to the HDI and PHDI modeling processes in East Java Province has been successfully carried out. Based on the analysis, it is concluded that the second-order bivariate binary logistic regression (S-BBLR) model produces a better classification accuracy than the first-order bivariate binary logistic regression model (F-

BBLR), with an accuracy of 86.84%. Modeling HDI and PHDI from Regencies/Cities in East Java Province with a S-BBLR model, the results obtained if all predictor variables have a significant effect either simultaneously or partially.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge financial support from the Institut Teknologi Sepuluh Nopember for this work, under project scheme of the Publication Writing and IPR Incentive Program (PPHKI).

CONFLICT OF INTERESTS

The authors declare that there is no conflict of interests.

REFERENCES

[1] V. Ratnasari, I.N. Budiantara, A.T.R. Dani, Nonparametric regression mixed estimators of truncated spline and gaussian kernel based on cross-validation (CV), generalized cross-validation (GCV), and unbiased risk (UBR) methods, Int. J. Adv. Sci. Eng. Inform. Technol. 11 (2021) 2400-2406. https://doi.org/10.18517/ijaseit.11.6.14464..

[2] A.T.R. Dani, V. Ratnasari, I.N. Budiantara, Optimal knots point and bandwidth selection in modeling mixed estimator nonparametric regression, IOP Conf. Ser.: Mater. Sci. Eng. 1115 (2021), 012020. https://doi.org/10.1088/1757-899X/1115/1/012020..

[3] D.C. Montgomery, E.A. Peck, G.G. Vining, Introduction to linear regression analysis, 5th ed., John Wiley & Sons, Inc., New Jersey, USA, 2012.

[4] K. Takezawa, Introduction to nonparametric regression, John Wiley & Sons, Inc., New Jersey, USA, 2006.

[5] J. Harezlak, D. Ruppert, M.P. Wand, Semiparametric regression with R, Springer, New York, 2018.

[6] D.W. Hosmer, S. Lemeshow, Applied logistic regression, 2nd ed., John Wiley & Sons, Inc., New Jersey, USA, 2000.

[7] Y.F. Gebrie, T.M. Dessie, N.F. Jemberie, Logistic regression analysis of exclusive breastfeeding practice among mothers in Amanuel Town, Northwest, Ethiopia, Amer. J. Data Mining Knowl. Disc. 3 (2018), 13-19.

[8] M.P. LaValley, Logistic regression, Circulation. 117 (2008), 2395–2399. https://doi.org/10.1161/CIRCULATIONAHA.106.682658.

[9] M.W. Fagerland, D.W. Hosmer, A generalized hosmer–lemeshow goodness-of-fit test for multinomial logistic regression models, Stata J. 12 (2012), 447–453. https://doi.org/10.1177/1536867X1201200307.

[10] H. Cui, L. Wu, S. Hu, R. Lu, S. Wang, Research on the driving forces of urban hot spots based on exploratory analysis and binary logistic regression model, Trans. GIS. 25 (2021), 1522–1541. https://doi.org/10.1111/tgis.12739.

[11] S. Mehrolia, S. Alagarsamy, V.M. Solaikutty, Customers response to online food delivery services during COVID-19 outbreak using binary logistic regression, Int. J. Consum. Stud. 45 (2021), 396-408. https://doi.org/10.1111/ijcs.12630.

[12] I.G. Sudarman, A. Ahmad, Mapping of landslide-prone areas in the Lisu river basin Barru Regency based on binary logistic regression, IOP Conf. Ser.: Earth Environ. Sci. 807 (2021), 022081. https://doi.org/10.1088/1755-1315/807/2/022081.

[13] Z. Cao, M.Y. Wong, Approximate maximum likelihood estimation for logistic regression with covariate measurement error, Biometrical J. 63 (2021), 27–45. https://doi.org/10.1002/bimj.202000024.

[14] M. Sabreen, D.K. Behera, Rural household's participation in non-farm economic activities in India using binary logistic regression model, Region. Sect. Econ. Stud. 21 (2021), 5-16.

[15] M. Fathurahman, Purhadi, Sutikno, V. Ratnasari, Hypothesis testing of geographically weighted bivariate logistic regression, J. Phys.: Conf. Ser. 1417 (2019), 012008. https://doi.org/10.1088/1742-6596/1417/1/012008.

[16] M.A. Heise, R.H. Myers, Optimal designs for bivariate logistic regression, Biometrics. 52 (1996), 613. https://doi.org/10.2307/2532900.

[17] B.W. McDonald, Estimating logistic regression parameters for bivariate binary data, J. R. Stat. Soc.: Ser. B (Methodol.) 55 (1993), 391–397. https://doi.org/10.1111/j.2517-6161.1993.tb01909.x.

[18] C.V. Ananth, J.S. Preisser, Bivariate logistic regression: modelling the association of small for gestational age births in twin gestations, Statist. Med. 18 (1999), 2011–2023. https://doi.org/10.1002/(SICI)1097-0258(19990815)18:15<2011::AID-SIM169>3.0.CO;2-8.

[19] K. Bel, D. Fok, R. Paap, Parameter estimation in multivariate logit models with many binary choices, Econ. Rev.

37 (2018), 534–550. https://doi.org/10.1080/07474938.2015.1093780.

[20] G.M. Fitzmaurice, N.M. Laird, G.E.P. Zahner, C. Daskalakis, Bivariate Logistic Regression Analysis of Childhood Psychopathology Ratings using Multiple Informants, Amer. J. Epidemiol. 142 (1995), 1194–1203. https://doi.org/10.1093/oxfordjournals.aje.a117578.

[21] L. Deldossi, S.A. Osmetti, C. Tommasi, Optimal design to discriminate between rival copula models for a bivariate binary response, TEST. 28 (2019), 147–165. https://doi.org/10.1007/s11749-018-0595-1.

[22] P. Purhadi, M. Fathurahman, A logit model for bivariate binary responses, Symmetry. 13 (2021), 326. https://doi.org/10.3390/sym13020326.

[23] M.A. Islam, R.I. Chowdhury, L. Briollais, A bivariate binary model for testing dependence in outcomes, Bull. Malaysian Math. Sci. Soc. 35 (2012), 845–858.

[24] M. Fathurahman, Purhadi, Sutikno, V. Ratnasari, Geographically weighted multivariate logistic regression model and its application, Abstr. Appl. Anal. 2020 (2020), 8353481. https://doi.org/10.1155/2020/8353481.

[25] E. Sofilda, P. Hermiyanti, M.Z. Hamzah, Determinant variable analysis of human development index in Indonesia (Case for high and low index at period 2004-2013), OIDA Int. J. Sustain. Dev. 8 (2015), 11–28. https://ssrn.com/abstract=2695465.

[26] Suparmi, N. Kusumawardani, D. Nambiar, Trihono, A.R. Hosseinpoor, Subnational regional inequality in the public health development index in Indonesia, Glob. Health Action. 11 (2018), 41–53. https://doi.org/10.1080/16549716.2018.1500133.

[27] E.K. Berndt, B.H. Hall, R.E. Hall, J.A. Hausman, Estimation and inference in nonlinear structural models, Ann. Econ. Soc. Measure. 3/4 (1974), 653-665.

[28] W.H. Greene, Econometric analysis, 5th ed., Pearson Education, Inc., New Jersey, USA, 2003.

[29] R.V. Hogg, A.T. Craig, Introduction to mathematical and statistics, 4th ed., Macmillan Publishing Co., Inc., New York, 1978.