J. Indones. Math. Soc.
Vol. 31, No. 03 (2025), pp. 1–9.

Semiparametric Bivariate Probit Model Early Breastfeeding Initiation and Exclusive Breastfeeding

Suci Amalia[1], Vita Ratnasari[2], Ismaini Zain[3]

[1,2,3]Department of Statistics, Institut Teknologi Sepuluh Nopember, Indonesia
[1]amaliasuci117@gmail.com, [2]vita_ratna@statistika.its.ac.id [3]ismaini_z@statistika.its.ac.id

Abstract. Regression analysis in which the response variable is categorical can be processed using the probit model. The probit model is based on the normal distribution, in addition to its interpretation based on marginal effect values. A probit model consisting of two response variables is called the bivariate probit model, in which the response variables each consist of two categories. The predictor variables in bivariate probit model can be either categorical and also continuous data. Bivariate probit model both response variables have a relationship. One of the developments of the bivariate probit model is the semiparametric bivariate probit model, where the bivariate probit model in which there is a parametric and a nonparametric model in this case in the form of a continuous covariate smooth function. Semiparametric bivariate probit model have the advantage of being able to address the problem of nonlinearity of undetected continuous predictor variables that can cause modeling inaccuracies that can effect the results of estimation accuracy. Parameter estimation of semiparametric bivariate probit model uses the Penalized Maximum Likelihood Estimation approach, but the equation obtained is not closed form so iteration are needed to solve it. The iteration used is Fisher Scoring. The semiparametric bivariate probit model was applied to data on early breastfeeding initiation and exclusive breastfeeding in East Java Province in 2021 with variables that affect early breastfeeding initiation being birth attendants while those affecting exclusive breastfeeding are maternal age and maternal education level.

Key words and Phrases: Early Breastfeeding Initiation, Exclusive Breastfeeding, Fisher Scoring, Penalized Maximum Likelihood Estimation, Semiparametric Bivariate Probit.

1. INTRODUCTION

A probit model consisting of two response variables is called the probit bivarit model, in which the response variables each consist of two categories. The predictor

2020 Mathematics Subject Classification: 62P10, 62J12, 62G08
Received: 25-03-2025, accepted: 27-05-2025.

variables in bivariate probit models can be both categorical and continuous, and the response variables has a relationship. In a probit model, the probability of an event is calculated using the cumulative distribution function of the standard normal distribution [1]. Many bivariate probit models have been developed, [2] developed a bivariate probit model with a semiparametric version and introduced a fitting procedure that allows to model binary result systems as a function of parametric terms and a continuous covariate smooth function, in other words a semiparametric bivariate probit model is a bivariate probit model in which there is a parametric model and a nonparamteric model which is a smooth function of continuous covariates. The smooth function in the semiparametric bivariate probit model uses a spline regression approach to anticipate nonlinearity [3]. Parameter estimation in bivariate probit models generally uses maximum likelihood but in semiparametric bivariate probit models parameter estimation uses penalized maximum likelihood estimation. [2] said the semiparametric bivariate probit model in the parameter estimation process can also use maximum likelihood, but to prevent overfitting, penalized maximum likelihood is used instead. In the parameter estimation process, the semiparameteric bivariate probit model produces equations that are not closed so that they require an iteration process, therefore the Fisher Scoring iteration method is used. Fisher Scoring method is a method developed from Newton Raphson in which the Hessian matrix is replaced with Fisher's information matrix. Iteration using fisher scoring is considered to have better performance in overcoming the possibility of non-convergence [4].

Early breastfeeding initiation is a breastfeeding process starting immediately after the baby is born, early breastfeeding initiation can be done by skin-to-skin contact between the baby and the mother that lasts for one hour. Early breastfeeding initiation has many benefits, namely being able to increase the success rate of breastfeeding, stimulate milk production, increase bonding between mother and baby to prevent infant death. Early breastfeeding initiation can also help the continuity of exclusive breastfeeding and the duration of breastfeeding. According to the Ministry of Health, exclusive breastfeeding is breastfeeding in the absence of food supplements or other additional drinks other than drugs. The benefits of exclusive breastfeeding in infants are not only as nutritional fulfillment but also help the inner bond between mother and baby and also increase intelligence in children. The World Health Organization recommends exclusive breastfeeding starting from the beginning of birth to six months of age and breastfeeding until the age of two with complementary foods. Early breastfeeding initiation and exclusive breastfeeding can be approached using a semiparametric bivariate probit model because both are categorical data where it is in accordance with the modeling requirements using a semiparametric bivariate probit model, namely the response variable is categorical and another condition, each response variable consists of two categories and is interrelated where in fact if early breastfeeding initiation is carried out then the potential for Exclusively breastfeeding is very large and also strengthened by the results of the independence test, namely early breastfeeding initiation and exclusive breastfeeding have a relationship.

2. METHODS AND DATA

This section discusses theories related to the semiparametric bivariate probit model

2.1. **Bivariate Probit Model.**

Bivariate probit models consisting of response variables Y_1 and Y_2 where both are unobserved response variables y_1^* and y_2^* with equations can generally be written in equations (1) and (2) [5].

$$y_1^* = \mathbf{x}^T \beta_1 + \varepsilon_1 \tag{1}$$

$$y_2^* = \mathbf{x}^T \beta_2 + \varepsilon_2 \tag{2}$$

where,

$$\begin{aligned} \mathbf{x} &= \begin{bmatrix} 1 & x_1 & \dots & x_k \end{bmatrix}^T, \text{ sized } (k+1) \times 1 \\ \beta_2 &= \begin{bmatrix} \beta_{20} & \beta_{21} & \dots & \beta_{2k} \end{bmatrix}^T, \text{ sized } (k+1) \times 1 \\ k &= \text{the multiplicity of predictor variables} \end{aligned}$$

2.2. **Semiparametric Bivariate Probit Model.**

The semiparametric bivariate probit model is a bivariate probit model that contains a parametric model and a nonparamteric model in this case a continuous covariate smooth function [6]. The semiparametric bivariate probit model has the advantage of being able to handle the problem of nonlinearity of undetected continuous predictor variables because it is able to cause modeling inaccuracies that affect the results of estimation accuracy [7]. Here is the equation of the semiparametric bivariate probit model.

$$y_{1i}^* = \mathbf{x}_{1i}^T \beta_1 + \sum_{p_1=1}^{P_1} f_{1p_1}\left(z_{1p_1 i}\right) + \varepsilon_{1i} \tag{3}$$

$$y_{2i}^* = \mathbf{x}_{2i}^T \beta_2 + \sum_{p_2=1}^{P_2} f_{2p_2}\left(z_{1p_2 i}\right) + \varepsilon_{2i} \tag{4}$$

y_1^* and y_2^* is a latent variable, with $i = 1, 2, \dots, n$ where n is the number of samples. To determine categories y_{1i} and y_{2i} based on criteria.

$$y_{wi} = \begin{cases} 0, y_{wi}^* \leq 0 \\ 1, y_{wi}^* > 0 \end{cases}, w = 1, 2 \tag{5}$$

$\mathbf{x}_{1i}^T = \begin{pmatrix} 1 & x_{12i} & \dots & x_{1p_1 i} \end{pmatrix}$ where i is the vector of $\mathbf{X}_1 = \begin{pmatrix} \mathbf{x}_{11} & \dots & \mathbf{x}_{1n}{}^T \end{pmatrix}^T$ the model matrix $n \times P_1$ containing parametric model components P_1 with corresponding parameter vectors β_1. $f_{1p_1 i}$ is an unknown smooth function of continuous covariates $z_{1p_1 i}$ and P_1 is the number of regressors $z_{1p_1 i}$. For identification purposes, smooth components are sensitive to constraints such as $\sum_i f_{wp_w}\left(z_{wp_w i}\right) = 0$.

The error $(\varepsilon_{1i}, \varepsilon_{2i})$ is assumed to follow a bivariate distribution.

$$\begin{pmatrix} \varepsilon_{1i} \\ \varepsilon_{2i} \end{pmatrix} \sim IIDN\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}\right) \tag{6}$$

With ρ is a correlation coefficient and error variance is normalized into a unity to identify parameters in the model. Errors are assumed to correlate with each other, so parameter estimation is done together and the semiparametric bivariate probit regression model is written as follows.

$$P\left(Y_{1i} = 1, Y_{2i} = 1 | x_{1i}, x_{2i}\right) = \Phi_2\left(\eta_{1i}, \eta_{2i}, \rho\right) \tag{7}$$

Φ_2 shows a normal bivariate distribution function with zero mean, variance and correlation ρ. Parameter estimation method of semiparameteric bivariate probit model using Penalized Maximum Likelihood Estimation (PMLE). The penalized likelihood function used is [8].

$$l_q\left(\beta^*\right) = \ln L\left(\beta^*\right) - \frac{1}{2}\theta^T \mathbf{S}_\lambda \theta \tag{8}$$

with,

$$\begin{aligned} \ln L\left(\beta^*\right) &= \ln\left(\prod_{i=1}^{n} P\left(Y_{11i} = y_{11i}, Y_{10i} = y_{10i}, Y_{01i} = y_{01i}\right)\right) \\ &= \sum_{i=1}^{n} \left[_{11i} \ln p_{11i} + y_{10i} \ln p_{10i} + y_{01i} \ln p_{01i} + \right. \\ &\left. y_{00} \ln\left(1 - p_{11i}^{y11i} - p_{10i}^{y10i} - p_{01i}^{y01i}\right)\right] \end{aligned} \tag{9}$$

Because the equation obtained is not closed form, the solution uses an iteration method, namely Fishser Scoring. In Fisher's Scoring method, a gradient vector $\mathbf{g}\left(\beta^*\right)$ is needed, which is the first derivative of the likelihood function, and a Fisher Information Matrix $\Im$ which is the second derivative of the likelihood function.

$$\mathbf{g}\left(\beta^*\right) = \begin{pmatrix} \frac{\partial l_q(\beta^*)}{\partial \beta_1} \\ \frac{\partial l_q(\beta^*)}{\partial \beta_1} \\ \frac{\partial l_q(\beta^*)}{\partial \theta_1} \\ \frac{\partial l_q(\beta^*)}{\partial \theta_2} \\ \frac{\partial l_q(\beta^*)}{\partial \rho} \end{pmatrix} \tag{10}$$

and,

$$\Im = -E\begin{pmatrix} \frac{\partial^2 l_q(\beta^*)}{\partial\beta_1\partial\beta_1^T} & \frac{\partial^2 l_q(\beta^*)}{\partial\beta_2\partial\beta_1^T} & \frac{\partial^2 l_q(\beta^*)}{\partial\theta_1\partial\beta_1^T} & \frac{\partial^2 l_q(\beta^*)}{\partial\theta_2\partial\beta_1^T} & \frac{\partial^2 l_q(\beta^*)}{\partial\rho\partial\beta_1^T} \\ \frac{\partial^2 l_q(\beta^*)}{\partial\beta_1\partial\beta_2^T} & \frac{\partial^2 l_q(\beta^*)}{\partial\beta_2\partial\beta_2^T} & \frac{\partial^2 l_q(\beta^*)}{\partial\theta_1\partial\beta_2^T} & \frac{\partial^2 l_q(\beta^*)}{\partial\theta_2\partial\beta_2^T} & \frac{\partial^2 l_q(\beta^*)}{\partial\rho\partial\beta_2^T} \\ \frac{\partial^2 l_q(\beta^*)}{\partial\beta_1\partial\theta_1^T} & \frac{\partial^2 l_q(\beta^*)}{\partial\beta_2\partial\theta_1^T} & \frac{\partial^2 l_q(\beta^*)}{\partial\theta_1\partial\theta_1^T} & \frac{\partial^2 l_q(\beta^*)}{\partial\theta_2\partial\theta_1^T} & \frac{\partial^2 l_q(\beta^*)}{\partial\rho\partial\theta_1^T} \\ \frac{\partial^2 l_q(\beta^*)}{\partial\beta_1\partial\theta_2^T} & \frac{\partial^2 l_q(\beta^*)}{\partial\beta_2\partial\theta_2^T} & \frac{\partial^2 l_q(\beta^*)}{\partial\theta_1\partial\theta_2^T} & \frac{\partial^2 l_q(\beta^*)}{\partial\theta_2\partial\theta_2^T} & \frac{\partial^2 l_q(\beta^*)}{\partial\rho\partial\theta_2^T} \\ \frac{\partial^2 l_q(\beta^*)}{\partial\beta_1\partial\rho^T} & \frac{\partial^2 l_q(\beta^*)}{\partial\beta_2\partial\rho^T} & \frac{\partial^2 l_q(\beta^*)}{\partial\theta_1\partial\rho^T} & \frac{\partial^2 l_q(\beta^*)}{\partial\theta_2\partial\rho^T} & \frac{\partial^2 l_q(\beta^*)}{\partial\rho\partial\rho^T} \end{pmatrix} \tag{11}$$

2.3. **Data.**

The semiparametric bivariate probit method will be applied to data on early breastfeeding initiation and exclusive breastfeeding in East Java in 2021 and factors suspected of influencing, where the data is obtained from a National Socioeconomic Survey (SUSENAS) conducted by Badan Pusat Statistik (BPS).Here are the variables used variables used (Table 1).

TABLE 1. Research Variable

Variable	Deskription/Information	Category
Y_1	Early Breastfeeding Initiation	0 : No
		1 : Yes
Y_2	Exclusive Breastfeeding	0 : No
		1 : Yes
Z_1	Mother's Age	-
Z_2	First Age of Marriage	-
X_1	Mother's Education Level	1 : No Diploma
		2 : Elementary/Junior
		3 : Senior
		4 : College
X_2	Mother's Employment Status	0 : Working
		1 : Not Working
X_3	Birth Attendant	0 : Health Workers
		1 : Not Health Worker

3. RESULTS AND DISCUSSION

The semiparametric bivariate probit model of the two response variables correlates with each other, therefore it is necessary to check it first using an independence test.

3.1. **Independence Test.**

The results of the independence test show that χ^2 of 18,076 is more than $\chi^2_{0,1;1}$ of 2,706 and is reinforced with a p-value of 0,000 less than α of 0.05. So it was decided to reject H_0, which means that there is a significant relationship between early breastfeeding initiation and exclusive breastfeeding.

3.2. **Semiparametric Bivariate Probit Model.**

Modeling data on early initiation of breastfeeding and exclusive breastfeeding using semiparametric bivariate probit model. Modeling data on early initiation of breastfeeding and exclusive breastfeeding using semiparametric bivariate probit model.

$$
\begin{aligned}
\widehat{y}_{1i}^{*} = {} & -0,46522 + 0,13130\ x_{12i} + 0,09661 x_{13i} + 0,06848 x_{14i} - 0,05452 x_{21i} + \\
& 0,64920 x_{31i} + 0,04317 z1_{k_1}^{1} - 0,00649 z1_{k_2}^{1} - 0,01601 z1_{k_3}^{1} - 0,00734 z1_{k_4}^{1} + \\
& 0,01710 z1_{k_5}^{1} + 0,00626 z1_{k_6}^{1} + 0,01607 z1_{k_7}^{1} - 0,01483 z1_{k_8}^{1} + 0,10334 z1_{k_9}^{1} - \\
& 0,01875 z1_{k_{10}}^{1} - 0,00000 z2_{k_1}^{1} + 0,00000 z2_{k_2}^{1} + 0,00000 z2_{k_3}^{1} + 0,00000 z2_{k_4}^{1} + \\
& 0,00000 z2_{k_5}^{1} + 0,00000 z2_{k_6}^{1} - 0,00000 z2_{k_7}^{1} - 0,01245 z2_{k_8}^{1} - \\
& 0,03412 z2_{k_9}^{1} - 0,02567 z2_{k_{10}}^{1} \\
\widehat{y}_{2i}^{*} = {} & -0,73726 + 0,30547 x_{12i} + 0,23719 x_{13i} + 0,27701 x_{14i} - 0,04681 x_{21i} + \\
& 0,42926 x_{31i} + -0,02123 z1_{k_1}^{1} - 0,00328 z1_{k_2}^{1} + 0,01216 z1_{k_3}^{1} - 0,00453 z1_{k_4}^{1} - \\
& 0,01292 z1_{k_5}^{1} - 0,00497 z1_{k_6}^{1} + -0,01312 z1_{k_7}^{1} + 0,10254 1_{k_8}^{1} + 0,08257 z1_{k_9}^{1} - \\
& 0,10925 z1_{k_{10}}^{1} - 0,41991 z2_{k_1}^{1} - 0,11436 z2_{k_2}^{1} - 0,10927 z2_{k_3}^{1} - 0,06293 z2_{k_4}^{1} - \\
& 0,09429 z2_{k_5}^{1} - 0,07607 z2_{k_6}^{1} + 0,08859 z2_{k_7}^{1} + 0,07489 z2_{k_8}^{1} - 0,04092 z2_{k_9}^{1} - \\
& 0,05184 z2_{k_{10}}^{1}
\end{aligned}
$$

Test the significance of parameters on the early breastfeeding initiation variable (Table 2).

TABLE 2. Parameter Significance early breastfeeding initiation

	Z	p-value	Decision
intercept	-1,860	0,06287	Reject H_0
X_{12}	0,918	0,35874	Fail to Reject H_0
X_{13}	0,658	0,51063	Fail to Reject H_0
X_{14}	0,437	0,66190	Fail to Reject H_0
X_{21}	-1,011	0,31202	Fail to Reject H_0
X_{31}	2,875	0,00404	Reject H_0
		p-value	Decision
$f(Z_1)$		0,207	Fail to Reject H_0
$f(Z_2)$		0,196	Fail to Reject H_0

Based on the table above, it is partially known that the variable that has a significant effect on early breastfeeding initiation with a significant level is birth support.

The significance test of smoothing parameters on early breastfeeding initiation showed that the maternal age and first age of marriage variables did not significantly affect the early breastfeeding initiation variables.

Test the significance of parameters on exclusive breastfeeding variables (Table 3).

TABLE 3. Parameter Significance exclusive breastfeeding

	Z	p-value	Decision
intercept	-2,901	0,00372	Reject H_0
X_{12}	2,115	0,03447	Reject H_0
X_{13}	1,600	0,10951	Fail to Reject H_0
X_{14}	1,759	0,07850	Reject H_0
X_{21}	-0,875	0,38156	Fail to Reject H_0
X_{31}	1,867	0,06189	Reject H_0
		p-value	Decision
$f(Z_1)$		1,05×10-5	Reject H_0
$f(Z_2)$		0,32	Fail to Reject H_0

Based on the table above, it is partially known that the variables that have a significant effect on exclusive breastfeeding with a significant level are the level of education of mothers and birth attendants.

The significance test of smoothing parameters on exclusive breastfeeding showed that maternal age variables had a significant effect on exclusive breastfeeding variables.

To interpret early breastfeeding initiation and exclusive breastfeeding, for example, there are individuals in households where in this study mothers aged 15-49 years and have babies aged 0-23 months with the last college education, the mother's employment status is work and at the time of childbirth assisted by medical personnel, then.

$$\begin{aligned}\widehat{y}^*_{1i} &= -0,46522 + 0,13130\,(0) + 0,09661\,(0) + 0,06848\,(1) - 0,05452\,(1) + 0,64920\,(1)\\ &= 0,19794\\ \widehat{y}^*_{2i} &= -0,73726 + 0,30547\,(0) + 0,23719\,(0) + 0,27701\,(1) - 0,04681\,(1) + 0,42926\,(1)\\ &= -0,0778\end{aligned}$$

So that the probability of each is obtained $\widehat{p}_{11}, \widehat{p}_{10}, \widehat{p}_{01}, \widehat{p}_{00}$

$$\begin{aligned}\widehat{p}_{11} &= \Phi_2\left(\widehat{\eta}_{1i};\widehat{\eta}_{2i};\widehat{\rho}\right) = 0,2712912\\ \widehat{p}_{10} &= \Phi\left(\widehat{\eta}_{1i}\right) - \Phi_2\left(\widehat{\eta}_{1i};\widehat{\eta}_{2i};\widehat{\rho}\right) = 0,3071628\\ \widehat{p}_{01} &= \Phi\left(\widehat{\eta}_{2i}\right) - \Phi_2\left(\widehat{\eta}_{1i};\widehat{\eta}_{2i};\widehat{\rho}\right) = 0,1977024\\ \widehat{p}_{00} &= 1 - \Phi\left(\widehat{\eta}_{1i}\right) - \Phi\left(\widehat{\eta}_{2i}\right) + \Phi_2\left(\widehat{\eta}_{1i};\widehat{\eta}_{2i};\widehat{\rho}\right) = 0,2238436\end{aligned}$$

Based on the probability value that can be known, individuals in the household have a probability of 0.2712912 who initiate early breastfeeding and provide exclusive breastfeeding.

The calculation of marginal effects in a semiparametric bivariate probit model is useful to find out how much influence the change in predictor variables has on response variables where other variables are assumed to be constant. The following is a calculation of marginal effects for each predictor variable, namely the education level of the mother and birth attendant.

$$\frac{\partial \widehat{p}_{11}}{\partial x_1} = \frac{\partial \Phi_2 \left(\widehat{\eta}_{1i}; \widehat{\eta}_{2i}; \widehat{\rho}\right)}{\partial x_1} = 0,1369666$$

From the calculation of the marginal effect for the variable level of maternal education is known to be 0.1369666, so it can be interpreted that individuals in the household, namely mothers with a college education level, are able to initiate early breastfeeding and provide exclusive breastfeeding by 13.69 percent themselves to mothers with non-college education levels.

$$\frac{\partial \widehat{p}_{11}}{\partial x_3} = \frac{\partial \Phi_2 \left(\widehat{\eta}_{1i}; \widehat{\eta}_{2i}; \widehat{\rho}\right)}{\partial x_3} = 0,4247015$$

From the calculation of the marginal effect for the birth attendant variable, it is known to be 0.4247015, so it can be interpreted that individuals in the household, namely mothers with childbirth assisted by medical officers, are able to initiate early breastfeeding and provide exclusive breastfeeding by 42.47 percent of themselves to mothers whose labor process is not assisted by medical workers (non-medical).

4. CONCLUSION

Data modeling of early breastfeeding initiation and exclusive breastfeeding using a semiparametric bivariate probit model found that variables that affect early breastfeeding initiation are birth attendants while those that affect exclusive breastfeeding are maternal age and maternal education level. Continuous variables, namely maternal age and first age of marriage, did not significantly affect the variable of early breastfeeding initiation, but for the variable of exclusive breastfeeding, it was known that only the age variable had a significant effect.

REFERENCES

[1] D. N. Gujarati, “Basic econometrics fourth edition,” *McGraw-Hill Education*, 2003.

[2] G. Marra and R. Radice, “Estimation of a semiparametric recursive bivariate probit model in the presence of endogeneity,” *Canadian Journal of Statistics*, vol. 39, no. 2, pp. 259–279, 2011. `https://doi.org/10.1002/cjs.10100`.

[3] G. Marra and R. Radice, “A penalized likelihood estimation approach to semiparametric sample selection binary response modeling,” *Electronic Journal of Statistics*, 2013. `https://doi.org/10.1214/13-EJS814`.

[4] U. Okeh and I. Oyeka, "Estimating the fisher's scoring matrix formula from logistic model," *Am. J. Theor. Appl. Stat*, vol. 2, pp. 221–227, 2013. https://doi.org/10.11648/j.ajtas.20130206.19.

[5] W. Greene, "Econometric analysis fifth edition prentice hall," 2002.

[6] G. Marra, "On p-values for semiparametric bivariate probit models," *Statistical Methodology*, vol. 10, no. 1, pp. 23–28, 2013. https://doi.org/10.1016/j.stamet.2012.05.003.

[7] G. Marra, R. Radice, and S. Missiroli, "Testing the hypothesis of absence of unobserved confounding in semiparametric bivariate probit models," *Computational Statistics*, vol. 29, pp. 715–741, 2014. https://doi.org/10.1007/s00180-013-0458-x.

[8] R. Radice, G. Marra, and M. Wojtyś, "Copula regression spline models for binary outcomes," *Statistics and Computing*, vol. 26, pp. 981–995, 2016. https://doi.org/10.1007/s11222-015-9581-6.