A Bayesian Logit-Normal Model in Small Area Estimation

E Sunandi$^{1,2}$, A Kurnia$^{2\ast}$, K Sadik$^2$ and K A Notodiputro$^2$

$^1$ Statistics Program, Department of Mathematics, Bengkulu University, Bengkulu, Indonesia
$^2$ Department of Statistics, IPB University, Bogor, Indonesia

$\ast$E-mail: anangk@apps.ipb.ac.id

Abstract. The Logit-Normal model is one of the GLM Bayes models with random covariates used in binary data. This study aimed to examine and evaluate the characteristics of the Logit-Normal model. The second objective was to apply the Logit-Normal model to estimate the proportion of poverty in the small area of Mukomuko District in Bengkulu Province. We used the Hierarchical Bayes (HB) method to estimate parameters model. The simulation results obtained from the optimum number of iterations and gibbs samples, namely 500 iterations and 100 gibbs samples, respectively. In addition, when seen the value of $\hat{p}_{i}^{HB}$ has the same tendency with the parameter $p_i$ and $\hat{p}_{i}^{DE}$. The value of $\hat{p}_{i}^{HB}$ tends to overestimate at $\hat{p}_{i}^{HB} \leq 60\%$. Conversely, $\hat{p}_{i}^{HB}$ tends to underestimate $\hat{p}_{i}^{HB} > 60\%$. Furthermore, in the simulation results the estimated value of variance and MSE $\hat{p}_{i}^{HB}$ tends to be smaller than the variance of proportion. The application of the HB method for Logit-Normal model in the Mukomuko District poverty data produces $\hat{p}_{i}^{HB}$ which has the same tendency as the result of direct estimator ($\hat{p}_{i}^{DE}$). The majority (21 villages) value of var($\hat{p}_{i}^{HB}$) are smaller or equal to the var($\hat{p}_{i}^{DE}$). This indicates that the estimation using the HB method can improve the estimation of the proportion parameters obtained by using the direct estimation method on poverty data.

1. Introduction

Regression analysis is often used to determine the relationship between the dependent variable and one or more independent variables. Regression analysis that is commonly used is classical regression analysis, where the dependent variable is continuous data that follows the Normal distribution. However, in its development the classical regression model was unable to overcome problems where the dependent variable was discrete data and was not normally distributed. The solution to overcome this problem is to use Generalized Linear Model (GLM). Gill and Torres [1] define that GLM is an extension of the classic linear model [1]. An important property of this model assumes the independence of observation. Normality and homogeneity of variance are not needed, so the relationship between dependent and independent variables with error distributions other than normal and heterogeneous variances can be modeled with this model.

There are two approaches in GLM modeling. They are the Bayes approach and the Frequentist approach. The Frequentist approach regards parameter values as a fixed effect. Whereas, the Bayes approach assumes that the data is fixed and that the parameters of the random model which are analyzed require the distribution of prior parameters. The Bayes approach involves two distributions, namely:
prior and posterior distribution. Bayes estimator is obtained from the expected value of posterior distribution [2].

In Bayesian GLM for binary data, such as logistics, it is usually not recommended to use a variety of prior distributions for $\beta$. One priors that can be used is independent $N(\mu, \sigma^2)$. This is because the posterior Bayes distribution estimator is well defined, even if the data have complete separation [2]. With the logit connecting function, Normal priors that are appropriate for linear predictors $x_i' \beta$ induce prior distributions for $P(y_1 = 1) = \exp(x'_i \beta)/(1 + \exp(x'_i \beta))$. This is a special case of the Logit-Normal distribution.

The Logit-Normal Model in Bayesian GLM has the assumption that the linear covariate component is fixed. But sometimes, the linear covariate component can be fixed and random. One alternative that can be used to help solve the Logit-Normal model is by using the Markov Chain Monte Carlo (MCMC) method which estimates the parameters of the population model using the estimation of the results of the Metropolis-Hasting algorithm.

One statistical method which is a mixed model is the model on Small Area Estimation (SAE). SAE combines the fixed effect of covariates and random effects of the area. SAE method is a statistical method for estimating parameters in a subpopulation which the number of examples is small or even non-existent. This method utilizes data from a large domain to infer variables that are of concern to smaller domains. The classical approach to estimating small area parameters is based on the application of the sampling design model and is known as the direct estimation method. The weakness of this method in subpopulations is that it does not have sufficient precision caused by the small number of samples used to obtain these allegations. The problems posed by this estimation method is the resulting estimator is not biased estimator but have a variant that is great and if on a small area to- i are not represented in the survey, it is not possible to do the estimation directly. Therefore, an indirect estimation method was developed. The purpose of this estimation method is to increase the effectiveness of sample size and reduce it to diversity so that it is more accurate. One method of estimating parameters used in SAE is Hierarchical Bayes (HB). This estimation method is used on discrete data that is count or binary. In this study, researchers are interested in reviewing the characteristics of the Logit-Normal model. This is due to data on factors/areas that affect poverty levels in Indonesia from a random process (Survey). Thus, the random effect of these factor areas cannot be ignored. In this study estimating the parameters of the model using the method of Hierarchical Bayes (HB) through Gibbs Sampling MCMC Metropolis-Hasting. Then, the model Logit-Normal applied to estimate proportion poverty in Area Small Mukomuko district.

2. Research Method

2.1 Exponential Family

The exponential family density function has several unique properties that are conducive to modeling. The random variable $Y$ is distributed as an exponential family. The canonical form can be written as follows [3]:

$$f(y|\eta) = \exp\left\{\frac{y\eta - c(\eta)}{a(\phi)} - d(y, \phi)\right\}$$

(1)

With $a(\cdot)$ is a function of the dispersion parameter $\phi$, $d(\cdot, \cdot)$ the data function and the dispersion parameter, and $c(\cdot)$ is a function of the parameter $\eta$ and is known as the cumulance function. The "linear" part of the mixed model stems from the fact that $\eta$ can be represented as a linear function of fixed and random parameters. $Y$ complete and stratified statistics; known as canonical statistics with corresponding canonical parameter $\eta$.

2.2 Bayesian Generalized Linear Model

The Bayes approach involves two distributions for $\theta$, namely: 1) Prior distribution $h(\theta)$, which is a function that states knowledge about $\theta$ before seeing the $y$ data ; and 2) Posterior distribution $h(\theta|y)$ is a function obtained from combining prior information with data information to update knowledge about parameters $\theta$ [2]. With Bayes theorem, the density function of posterior probability $h(\theta|y)$ is...
When observing $y$ and state $f(y|\theta)$ is a function of $\theta$, $f(y|\theta) = \ell(\theta)$, $\ell(\theta)$ is the likelihood function and $f(y)$ is a function of marginal probability $y$.

The main part of the Bayes theorem is the numerator $h(\theta|y) \propto f(y|\theta)h(\theta) = \ell(\theta)h(\theta)$, Bayes inference is dependent on data only through the likelihood function. For further observation $y^*$ posterior predictive distribution is a conditional distribution of $(y^*|y_1, ..., y_n)$. Probability density function of observation is $f(y^*|y) = \int f(y^*|y, \theta)h(\theta|y)d\theta$, that is, taking the spread to $y^*$ as if $\theta$ is known and then integrate with respect to the function of the concentration of probability posterior.

In Bayesian GLM, usually posterior distribution has no closed form expression. The difficulty is evaluating the integral of the denominator that determines $f(y)$, i.e. determining the appropriate constants so that the posterior is integrated into 1. The simulation method can approach the posterior distribution. The main method for doing it is the Markov Chain Monte Carlo (MCMC). The main MCMC methods are Gibbs sampling and the Metropolis-Hastings algorithm. Each requires that the function be proportional to the posterior distribution, for Bayesian infections to be correct after multiplying the prior distribution by the probability function.

Research on Bayesian GLM and Binary GLM has been carried out by several researchers, including: Thomson, et. al [4] performed simultaneous parameter stimulation and variable selection through continuous analogue Logit-Normal from prior spikes and slabs [4]; Molenberghs et al [5] conducted a hierarchical modeling with informal and random effects conjugates [5]; Hyunguk and Carl [2017] examine the Posterior Dependence of Data from Bayesian Beta-Binomial-Logit Mode [6]; Joyee, Yingbo, and Robin (2018) examined the use of Prior Cauchy distribution for Bayesian Logistics Regression [7]; Chong and David [2018] reviewed the General Method for the Solid Bayesian Modeling [8]; Xulong (2018) discusses the Bayesian framework for general linear mixture modeling identifying new loci candidates for late-onset Alzheimer’s disease [9]; Ogunsakin and Siaka (2019) reviewed the Bayesian Generalized Linear Modeling of Breast Cancer Blends [10]; and Dietz and Chatterjee [11] developed a Logit-Normal mixed model for Indian monsoon rainfall [11].

While the researchers who conducted the study Bayesian GLM at SAE are Franco C and Bell WR [12] review of information borrowed during over time on the Binomial models/Logit-Normal to SAE [12]; Torabi M and Shokoohi [13] study the estimation of small areas using the hierarchical bayes method used in cross-section data and time series [13]; Trevisani M and Torelli N [14] examine the comparison of the Hierarchical Bayes model for data counting on Estimating Small Areas [14]; Handayani et al [15] conducted the spatial empirical Bayes predictor of the small area mean for a lognormal variable of interest and spatially correlated random effects [15] and Clement [16] estimated small areas that were applied to disease spread data [16].

2.3 Model Logit Normal

The Normal Logit Model is an probability model with a hierarchical structure. This model has 2 levels of function. The second level is the random influence of an unobservable covariate. While the first level is a function of the density of opportunities of a random variable that belongs to an exponential family. Assumptions in the mixed model i.e observed data are independent of random effects. The Logit-Normal model has been used by previous researchers. Satopää et al [16] examines the Logit-Normal model to predict probability [17]. This model is used to build a probability-based aggregation formula to combine several probability estimates. Dietz dan Chatterjee [11] modeled rainfall Munson India using Logit-Normal model [11].

Then, the model can be written as follows:

**Level 1:**

$$Y_i|u \sim f(y|u) = \exp \left\{ \frac{\eta_i - c(\eta_i)}{a(\phi)} - d(y, \phi) \right\}$$

$$\eta_i = x_i^\prime \beta + z_i^\prime u$$

(2)

(3)
Level 2:

\[ \mathbf{u} \sim N_m(0, \Sigma) \]  

The \( \mathbf{\beta} \) vector is called a fixed effect. Random effects covariate \( \Sigma \) is a function of \( q \)-dimensional known (\( \sigma_1, \ldots, \sigma_q \)) as komponen variants. Whereas the px1 vector \( \mathbf{x}'_{i} \) is a fixed covariate. Random covariate vectors for \( i \)-th and \( r \)-data can be denoted by \( m_r \times 1 \) vector \( z_{ir} \). Then a vector is combined for each variant component to form a random covariate vector \( \mathbf{z}_i = (\mathbf{z}'_{i1}, \ldots, \mathbf{z}'_{iq})' \) with a length \( M = \sum_{r=1}^{q} m_r \). The random effect vector \( \mathbf{U} \) follows the \( m \)-dimensional Normal distribution with an average vector of \( \mathbf{0} \) and kovarianz \( \Sigma \).

The Logit-Normal model is formulated as follows [18]:

Level 1:

\[ Y_i \mid \mathbf{u} \sim \text{Binomial}(n_i, p_i) \]
\[ \text{logit}(p_i) = \eta_i = \mathbf{x}'_{i} \mathbf{\beta} + u_i \]  

Level 2:

\[ u_i \sim N(0, \sigma_u^2); i = 1, \ldots, m \]  

where \( \mathbf{\beta} \) and \( \sigma_u^2 \) are independent. Besides that, \( f(\mathbf{\beta}) \propto 1 \) and \( \sigma_u^{-2} \sim \text{Gamma}(a, b); a \geq 0, b > 0 \).

2.4 Markov Chain Monte Carlo Simulation (MCMC)

When using the Hierarchical Bayes (HB) method, posterior distribution calculations are required. This calculation is usually via multidimensional integrals. An alternative that can be used is to calculate the posterior magnitude through numerical integration. One of these methods is the Markov Chain Monte Carlo (MCMC). The main idea of the MCMC is to build a Markov chain probability in the end towards a certain posterior distribution. Calculation of posterior distribution results in samples of posterior quantities. Finally, parameters of posterior distribution can be estimated.

The famous MCMC procedure is conditional Gibbs. In the Logit-Normal, gibbs conditional model as follows [18]:

i. \( [\mathbf{\beta}, p, \sigma_u^2, y] \sim N_m \left[ \mathbf{\beta}^*, \sigma_u^2 \left( \sum_{i=1}^{N} X_i X'_i \right)^{-1} \right] \)

ii. \( [\sigma_u^2, p, \mathbf{\beta}, y] \sim \text{Gamma} \left[ \frac{N}{2} + \frac{1}{2} \left( (\mathbf{\eta} - X' \mathbf{\beta})' (\mathbf{\eta} - X' \mathbf{\beta}) + b \right) \right] \)  

iii. \( [p, \mathbf{\beta}, \sigma_u^2, y] \sim h(p, \mathbf{\beta}, \sigma_u^2) k(p) \)

Estimation of parameters \( \mathbf{\beta} \) and \( \sigma_u^2 \) raised directly from (i) and (ii). The parameter \( \mathbf{\beta}^* \) in part (i) of equation (9) is stated by \( \mathbf{\beta}^* = (\sum_{i=1}^{m} x_i x'_i)^{-1} (\sum_{i=1}^{m} x'_i \eta_i) \). Meanwhile, part (iii) Equation (9) is stated as follows:

a) \( f(p, \mathbf{\beta}, \sigma_u^2, y) \propto h(p, \mathbf{\beta}, \sigma_u^2) k(p) \)

b) \( h(p, \mathbf{\beta}, \sigma_u^2) = \frac{\partial \eta_i}{\partial p_i} \exp \left\{ \frac{-(n_i - x'_i \beta)^2}{2 \sigma_u^2} \right\} \)  

c) \( k(p) = p_1^y (1 - p_1)^{n_i - y_i} \)

The proportion value of HB will be estimated through the Gibbs Sampling Metropolis-Hasting (MH) simulation. The MCMC gibbs sample can be generated directly from (c) in Equation (10). The MH algorithm is as follows:

1. Took any value \( p_i^* \) from uniform distribution (0,1).
2. Generated \( \mathbf{\eta} \sim \text{ind N}(X' \mathbf{\beta}, \sigma_u^2 \mathbf{I}) \), then searched for value \( p_i^{(0)} = g^{-1}(\eta_i) \)
3. Counted \( r \left( p_i^{(d)}, p_i^* \right) = \min \left\{ k(p_i^{(d)}) / k(p_i^*), 1 \right\}; d = 0, 1, \ldots, D \)
4. Generated \( u \) from a uniform distribution (0,1).
5. Selected \( p_i^{(d+1)} = p_i^* \) if \( u \leq r \left( p_i^{(d)}, p_i^* \right) \).
6. Repeated step 3 until a \( D \) sample is obtained.
After performing an MH simulation, the following sequence of proportion estimators is obtained
\[ \{p_1^{(d)}, \ldots, p_m^{(d)}; d = 1, \ldots, D\}. \] Then the posterior quantities being observed can be calculated. Estimating the proportion of Hierarchical Bayes \( \hat{p}_{i i}^{\text{HB}} \) is
\[ \hat{p}_{i}^{\text{HB}} \approx \frac{1}{D} \sum_{d=1}^{D} p_i^{(d)} = p_i^{(i)}. \] Whereas the variance estimator of the proportion of HB is
\[ V(\hat{p}_{i}^{\text{HB}} | \hat{p}) \approx \frac{1}{D - 1} \sum_{d=1}^{D} (p_i^{(d)} - \hat{p}_i^{(i)})^2. \] (11)

Statistical properties of \( \hat{p}_{i}^{\text{HB}} \) which would allegedly namely Mean Square Error (MSE) and bias. MSE is obtained by taking the difference between the proportion of actual and proportion of HB for each area- \( i \), then squaring the difference and summing all areas. Squaring the differences each give weight greater for errors big. This amount is then divided by the number of areas to give MSE. While the bias of the proportion of HB is obtained from the difference between the actual proportion and the HB proportion for each area, then sums it all up and divides by the number of areas. The formula mathematically written as follows:
\[ \text{Bias} (\hat{p}_{i}^{\text{HB}}) = \left| \frac{1}{D} \sum_{d=1}^{D} \left[ p_i^{(d)} - p_i \right] \right| \] (12)
\[ \text{MSE}(\hat{p}_{i}^{\text{HB}}) = \left( \text{Bias} (\hat{p}_{i}^{\text{HB}}) \right)^2 + \text{Var}(\hat{p}_{i}^{\text{HB}}) \] (13)

As a measure of the accuracy and validation of estimating the proportion of HB in use Root Mean Square Error (RMSE) and the percentage of bias relative (Rbias). RMSE is the root of MSE. While Rbias can be interpreted as rasio biased estimator and its parameters. So, the most accurate estimate will lead to the smallest Rbias value going to zero. While the percentage of relative bias for each area is obtained from
\[ \text{Rbias}(\hat{p}_{i}^{\text{HB}}) = \frac{\text{Bias} (\hat{p}_{i}^{\text{HB}})}{p_i} \times 100\% \] (14)
\[ \text{RMSE} = \sqrt{\text{MSE}(\hat{p}_{i}^{\text{HB}})} \] (15)

### 2.5 Data
The data used are simulation data and secondary data. Simulation data is used to determine various characteristics of the Logit-Normal model. The simulation scenario is as follows:
**Table 1. Simulation Scenario**

| Item | Value |
|------|-------|
| $p_i$ | 0.021 0.021 0.030 0.041 0.044 0.060 0.077 0.073 0.089 0.118 0.146 0.191 |
|      | 0.238 0.322 0.337 0.384 0.459 0.493 0.560 0.584 0.673 0.727 0.758 0.823 |
|      | 0.856 0.914 0.921 0.934 0.943 0.941 0.944 0.956 0.980 0.983 0.996 |
| $N_i$ | 938 1178 847 962 1299 1152 1416 1294 1440 1182 1158 1299 |
|      | 952 1143 1163 825 850 1185 1376 977 1173 1343 1311 1088 |
|      | 1144 880 945 1267 1489 1460 975 897 1173 1482 927 |
| $n_i$ | 19 47 17 10 52 12 71 13 14 24 35 26 |
|      | 48 46 58 33 9 36 14 29 35 54 39 33 |
|      | 46 44 47 25 74 15 39 45 12 15 28 |
| $y_i$ | 1 1 2 1 3 1 5 1 1 2 5 5 |
|      | 12 14 20 13 4 18 8 18 23 38 29 26 |
|      | 39 40 43 23 69 14 37 43 11 15 28 |
| Iteration | 5 10 50 100 500 1000 |
| Sampel Gibbs | 5 10 50 100 500 1000 |

**number of area (m) = 35  a = 0.05  b = 0.0005**

While the empiric data obtained from the Statistic Indonesia in 2015. As a unit observation is the village in Mukomuko District Bengkulu Province. The variables used in this study are dependent variables, namely the proportion of the poor population. Whereas the independent variable is the percentage of illiterate women, the percentage of families without access to electricity, the number of health facilities, the number of malnutrition sufferers in the last 3 years, the number of maternal deaths during childbirth, the number of educational facilities, the number of recipients of the Community Health Insurance card (JAMKESMAS), the number of Certificates Not Able (SKTM). Some variables also had used in previous paper [19].
Table 2. Operational Definitions

| No. | Variable name                                      | Definition                                                                                                                                                                                                 | Units of measurement |
|-----|---------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------|
| 1   | Proportion of Poor Population (pi )               | The proportion of the poor is the proportion of the population below the poverty line. The cause of poverty is usually due to the scarcity of means to fulfill basic needs, or the difficulty of access to the fulfillment of work and education needs. | -                    |
| 2   | Percentage of illiterate women (X1)               | Percentage of women aged 15-44 years who cannot read and write.                                                                                                                                              | Percent              |
| 3   | Percentage of Families Without Electricity Access (X2) | Percentage of families that do not use electricity either PLN or non-PLN electricity.                                                                                                                                 | Percent              |
| 4   | Number of Health Facilities (X3)                  | Health facilities include hospitals, maternity hospitals/maternity hospitals, polyclinics/clinics, puskesmas, assisting puskesmas, doctor's practice, midwife's practice, poskesdes (village health post), polindes (village maternity hut), and posyandu located at every village | Unit                 |
| 5   | Number of Sufferers of Malnutrition in the Last 3 Years (X4) | Malnutrition is a form of malnutrition. Basically, poor nutrition can be interpreted as a condition where a person lacks an intake that contains energy and protein. | Person               |
| 6   | Number of Maternal Deaths During Childbirth (X5)  | Number of mothers giving birth who die.                                                                                                                                                                    | Person               |
| 7   | Number of Educational Facilities (X6)             | Educational facilities used are formal education including elementary school/ equivalent, junior high school / equivalent, high school/equivalent, and higher education (Diploma and undergraduate). | Unit                 |
| 8   | Number of Recipients of Community Health Insurance Cards (JAMKESMAS) (X7) | Community Health Insurance (J Amkesmas ) is a social assistance program for health services for the poor and disadvantaged that is carried out nationally, in order to realize comprehensive health services for the poor | Person               |
| 9   | Number of Certificate of Disability (SKTM) (X8)   | Certificate of Disability (SKTM) is a letter issued by the Village for Poor Families (Gakin ) to get convenience in their lives both in the social, health, economic and educational fields.                                    | Unit                 |

2.6 Method
The research method was carried out in two stages, the first evaluating the characteristics of the Logit-Normal model and the second the application of the Logit-Normal model to poverty data. The simulation algorithm is as follows:
1. Generated vector X is from normal multivariate distribution with specified parameters.
2. Determined $p_i$, $N_i$, $n_i$, $y_i$, $m$, $a$, $b$
3. Counted \( \hat{p}_i^{DE} = \sum_{i=1}^{m} \frac{y_i}{n_i} \)
4. Determined \( \beta \)
5. Generated \( \sigma_0^2 \sim \text{IG}(a, b) \)
6. Formed \( M = \text{diag}(\sigma_0^2, m) \)
7. The simulation through Metropolis-Hasting:
   a. Took any value \( p_i^{*} \) from uniform distribution \((0, 1)\).
   b. Generated \( \eta \sim \text{ind} N(X \beta, M) \), then searched for value \( p_i^{(0)} = g^{-1}(\eta_i) \)
   c. Counted \( r \left( p_i^{(d)}, p_i^{*} \right) = \min \left\{ \frac{k(p_i^{*})}{k(p_i^{(d)})}, 1 \right\} ; d = 0, 1, ..., D \)
   d. Generated \( u \) from a uniform distribution \((0, 1)\).
   e. Selected \( p_i^{(d+1)} = p_i^{*} \) if \( u \leq r \left( p_i^{(d)}, p_i^{*} \right) \).
   f. Repeat step c) until e) D sample is obtained.
8. Calculated HB-proportions, HB-proportion variance, HB-proportion bias, and HB-HB proportions
9. Calculated variance and MSE \( \hat{p}_i^{DE} \). The variant parameter formula is as follows:
   \[
   \text{var} \left( \hat{p}_i^{DE} \right) = \frac{(N_i - n_i)}{(N_i - 1)} \frac{\hat{p}_i^{DE} (1 - \hat{p}_i^{DE})}{n_i - 1}
   \]
10. Calculated relative bias and RMSE
11. Drew a conclusion.

Procedure for analyzing poverty data using the Logit-Normal model is
1. Preparation of dependent variables
   a. Calculate per capita expenditure per household. Data obtained from the 2015 Statistics Bengkulu Province data.
   b. Classify poor families (1) and non-poor (0) based on the Poverty Line. If a household has a per capita expenditure below the GK then poor families are classified (1), if vice versa (0).
   c. Make a direct estimate of the proportion of poor families in each of the surveyed villages:
   \[
   \hat{p}_i^{DE} = \sum_{i=1}^{m} \frac{y_i}{n_i}
   \]
2. Preparation of auxiliary variables
   a. Define the auxiliary variables to be used in constructing the model and estimating the HB estimator. The variables were taken from from the census.
3. Estimating the proportion of Hierarchical Bayes
   a. Estimated \( \beta \) with generate equation (9) section (i).
   b. Estimated \( \sigma_0^2 \) with generate equation (9) section (ii).
   c. Estimated the posterior distribution of the Hierarchical Bayes equation (10) part (iii) by entering parameters \( p_i^{DE} \).
   d. Perform a Metropolis-Hasting (MH) simulation through equation (10).
   e. Estimating the proportion of Hierarchical Bayes
   f. Predicted variance of HB estimator
4. Compared variance of direct estimator and HB estimator
5. Drew a conclusion.

3. Results and Discussion
3.1 Characteristics of The Logit-Normal Model
The simulation is designed to study and evaluate the statistical characteristics of the Logit-Normal model estimator parameters. Simulations carried out by a combination of the number of iterations and gibbs sampling at 5, 10, 50, 100, 500, and 1000. The simulation results the average RMSE\( \left( \hat{p}_i^{HB} \right) \) are presented in Table 3. The table can be seen that the value RMSE\( \left( \hat{p}_i^{DE} \right) \) for each iteration decreases with
increasing number of gibbs sample. The smallest RMSE(\hat{p}_{i\text{HB}}) value is obtained when the number of samples over 50 gibbs and iteration over 100 times by 0.036. This can be used as a reference that modeling using the Logit-Normal model in these conditions has good performance. The author chooses the number of gibbs 500 samples and 100 iterations to be stable (optimum) conditions in the modeling. This is because the more iterations and samples of gibbs it takes longer computation time.

Table 3. Average RMSE

| Iteration | 5     | 10    | 50    | 100   | 500   | 1000   |
|-----------|-------|-------|-------|-------|-------|--------|
| Sample    | Gibbs |       |       |       |       |        |
| 5         | 0.072 | 0.052 | 0.043 | 0.039 | 0.037 | 0.037  |
| 10        | 0.055 | 0.044 | 0.040 | 0.039 | 0.037 | 0.036  |
| 50        | 0.038 | 0.039 | 0.037 | 0.037 | 0.036 | 0.036  |
| 100       | 0.037 | 0.038 | 0.037 | 0.036 | 0.036 | 0.036  |
| 500       | 0.036 | 0.036 | 0.036 | 0.036 | 0.036 | 0.036  |
| 1000      | 0.036 | 0.036 | 0.036 | 0.036 | 0.036 | 0.036  |

Dietz and Chatterjee [11] conducted a simulation with four GLMM algorithms which is a method of estimating the parameters of the Normal Logit model. The four parameter estimation methods are Penalized Quasi Likelihood (PQL), The Penalized Iteratively Rewarded Least Squares (PIRLS), Method of Simulated Moments (MSIM), and The data cloning to be used in conjunction with MCMC. One result shows that estimating parameters using MCMC is good but requires a long computational time [11]. This is in line with the simulations conducted by the author.

Figure 1. Comparation of parameter \(p_i\), \(\hat{p}_i^{DE}\), and \(\hat{p}_i^{HB}\) at iteration 500 and Gibbs samples 100

Pattern comparison between parameters \(p_i\), \(\hat{p}_i^{DE}\), and \(\hat{p}_i^{HB}\) when the number of iterations and samples Gibbs respectively 500 and 100 can be seen in Figure 1. Third values have the same trend pattern, especially in [0.1,0.9]. This reflects that the formed model has good flexibility. It can be seen from the alleged plot that is able to follow the pattern of distribution of proportion parameters. However, at \(\hat{p}_i^{HB} \leq 0.6\) appears that the proportion of HB estimators overestimate. Whereas on \(\hat{p}_i^{HB} > 0.6\), then the value dugaanya estimate the proportion of use-Normal logit models have estimators underestimate. The largest difference in parameter values and the HB-estimator occurs when \(\hat{p}_i^{HB} \leq 0.1\) and \(\hat{p}_i^{HB} > 0.9\).

Meanwhile, Bias \(\hat{p}_i^{DE}\) and \(\hat{p}_i^{HB}\) relatively small (Figure 2 (a)). However, the bias \(\hat{p}_i^{HB}\) is greater than that \(\hat{p}_i^{DE}\). It is undeniable that \(\hat{p}_i^{HB}\) is a slightly biased estimator. But the alleged bias \(\hat{p}_i^{HB}\) relatively small is bias(\(\hat{p}_i^{HB}\)) \leq 1.3%. Furthermore, when seen in Figure 2 (b), the \(Var(\hat{p}_i^{HB})\) is much smaller than \(Var(\hat{p}_i^{DB})\). This is in accordance with the advantages of the SAE model that can reduce the estimator variance. This has implications for the MSE and RMSE values of
proportion estimators. In Figure 2 (c) it appears that the $RMSE(\hat{p}_i^{DE})$ is greater than $RMSE(\hat{p}_i^{HB})$. However at interval $\hat{p}_i^{HB} < 10\%$ and $\hat{p}_i^{HB} > 90\%$, the value of MSE is almost the same as MSE. This is because the bias in that range is large. This situation shows that the normal Logit model is poor that is used to estimate pi values that are relatively small ($\hat{p}_i^{HB} < 10\%$) or relatively large ($\hat{p}_i^{HB} > 90\%$). These conditions can be improved by expanding Logit-Normal model to become a model bivariate and time series to reduce the variance of the prediction error of the Logit-Normal model univariate on the value of the alleged pi relatively small ($\hat{p}_i^{HB} < 10\%$) or relatively large ($\hat{p}_i^{HB} > 90\%$) [12]. But conversely, if the value $\hat{p}_i^{HB}$ is in the interval [0.1; 0.9], then the Logit-Normal model has the best performance.

On the other hand, if we consider the relative bias percentage value (Rbias) in Figure 2 (d), it appears that at 0.1 $\leq \hat{p}_i^{HB} \leq$ 0.9 the value has a relatively small value with increasing proportions. However, the Median Rbias($\hat{p}_i^{HB}$) is greater than Rbias($\hat{p}_i^{DE}$) (Figure 2 (d)). Rbias($\hat{p}_i^{HB}$) values extremized when relatively small ($\hat{p}_i^{HB} < 0.1$).

3.2 Data Description

In this study using random sample data from the BPS census block. There were 39 villages as samples. Next, Table 4 shows a summary of the statistical descriptions of the variables used. In this table, it can be seen that the average proportion of poor families in Mukomuko Regency is 0.14 with a standard balance of 0.13. The largest proportion value is 0.41 owned by the Banjar Sari Village. On the contrary, the lowest proportion of poor population is owned by Koto Jaya Village, Air Dikit Village, Lubuk Pinang Village, and Ujung Padang Village with a value of 0.01.
Table 4. Summary of statistical descriptions of the variables used

| Variable                                      | Average | Standard Deviation | Minimum | Maximum |
|-----------------------------------------------|---------|--------------------|---------|---------|
| Proportion of Poor People (pi)                | 0.14    | 0.13               | 0.01    | 0.41    |
| Percentage of illiterate women (X1)           | 6.79    | 5.93               | 0.00    | 25.13   |
| Percentage of Families Without Electricity Access (X2) | 0.03    | 0.06               | 0.00    | 0.23    |
| Number of Health Facilities (X3)              | 4       | -                  | 1.00    | 1:00 p.m.
| Number of Sufferers of Malnutrition in the Last 3 Years (X4) | -       | 1.41               | 0.00    | 8.00    |
| Number of Maternal Deaths During Childbirth (X5) | 0.69    | 1.10               | 0.00    | 5.00    |
| Number of Educational Facilities (X6)         | 4       | 3.02               | 1.00    | 4:00 p.m.
| Number of Recipients of Community Health Insurance Cards (JAMKESMAS) (X7) | 253     | 169.02             | 0.00    | 598.00  |
| Number of Certificate of Disability (SKTM) (X8) | 51      | 49.30              | 2.00    | 213.00  |

The next variable is the percentage of illiterate women (X1). In 2015, Pondok Kandang Village did not have illiterate women. While the highest percentage of illiterate women (X1) that owned by Dusun Pulau village is 25.13. Meanwhile, the largest percentage of families without access to electricity (X2) is owned by Pelokan Baru Village. Instead some villages have used access to electricity.

Furthermore, Bandar Ratu Village has the highest number of health facilities, which is 13 units. While some villages have only 1 health facility. Then, the majority of villages in Mukomuko Regency do not have the number of sufferers of malnutrition in the last 3 years (X4), even though the Rawa Bangun village has 8 sufferers of malnutrition in the last 3 years. Then, in the variable of the Number of Maternal Deaths During Childbirth, the Village of Attractor was ranked top of 5 people. While some villages do not have cases of mothers dying during childbirth.

In addition, Bumi Mekar Jaya Village has the most educational facilities, 16 units. While there are some villages that only have 1 unit of educational facilities. The number of JAMKESMAS recipients in Mukomuko District is very variant. In fact, Bumi Mekar Jaya Village does not have residents who receive JAMKESMAS. However, the village has the highest number of SKTM recipients (213). While the Banjar Sari and Dusun Pulau Village only issued 2 SKTMs.

3.3 Logit-Normal Model on Poverty Data

Mukomuko Regency consists of 152 villages. 25.66% of this number or 39 villages are the BPS census block, with the number of households for each village selected as examples ranging from 8 to 20 households. The number of samples for each village is very small compared to the number of households in each village, which only ranges from 0.1% to 1.26%. The small number of samples used directly causes MSE of a proportion direct estimator produced very large. To improve the results of the estimation of poor population pioneering, the Logit-Normal model has been used using HB method.
Figure 3. Comparison of Proportion estimator of Poor Population in Mukomuko District
Gibbs sample generation through the Metropolis algorithm is used with optimal iteration and gibbs sample 500 and 100, respectively. Estimates $\hat{p}_i^{HB}$ can estimate the proportion of poverty at the villages in Mukomuko District. Based on the estimated values in Table 5, it shows that the direct estimator of proportion $\left(\hat{p}_i^{DE}\right)$ of poor village in Mukomuko relatively diverse with values ranging between 1% to 41.0%. Meanwhile, the $\hat{p}_i^{HB}$ value ranging from 3.3% to 45.3%. In accordance to the simulation results that if $\hat{p}_i^{HB} < 60\%$, then all of $\hat{p}_i^{HB}$ tend to overestimate.

| Statistic          | Min  | Max  | Median | Mean  |
|--------------------|------|------|--------|-------|
| $\hat{p}_i^{DE}$   | 0.010| 0.410| 0.090  | 0.140 |
| $\hat{p}_i^{HB}$   | 0.033| 0.453| 0.178  | 0.211 |
| $\text{Var}(\hat{p}_i^{DE})$ | 0.001| 0.034| 0.009  | 0.012 |
| $\text{Var}(\hat{p}_i^{HB})$ | 0.003| 0.031| 0.012  | 0.013 |

Table 5. Summary Statistics Estimating the proportion of Hierarchical Bayes

The use of the HB method in the Logit-Normal model can produce HB-proportion values that are different from the results of the direct estimation ($p_i^{DE}$). However, if seen in Figure 3, some of the estimators have the same tendency. The model has a flexibility that is quite good, can be seen from the results of the alleged plot were able to follow the pattern of distribution of observation data. From these results, there are 3 villages that have a proportion of poor population more than 40%. They are Kota Praja Village, Banjar Sari Village, and Agung Jaya Village. The smallest poverty proportion was in Pondok Kandang Village. While the largest proportion of HB estimators was in the Kota Praja Village.

Some villages have the value of HB estimator variance greater than variance estimators of direct proportion (Figure 4). This is consistent with the simulation results that the Logit-Normal model is inappropriate that is used to estimate pi values that are relatively small ($\hat{p}_i^{HB} < 10\%$) or relatively large ($\hat{p}_i^{HB} > 90\%$). In Figure 4, it can be seen that the majority (21 villages) of the variance value of the HB-proportion estimator have a value that is smaller or equal to the varian direct estimator of the proportion of pi. Therefore, it can be claimed that estimation by the HB method can improve the estimation of the proportion parameters obtained by using the direct estimation method.
4. Conclusion

From simulation results can be seen that with increasing number of iterations and samples gibbs then the average RMSE of HB estimator shrink until constant at a value 0.036. Simulation results can also be obtained by the number of iterations and samples gibbs optimum 500 iterations and 100 sample gibbs, respectively. Moreover, if seen the values of HB estimator have trand the same parameters. Values of HB estimator tend to overestimate that is greater than 60% and tend to underestimate that is less than 60%. Furthermore, the results of the simulation predicted HB MSE is smaller than the MSE of direct estimation proportion for each area. If the proportion is in the interval [0.1; 0.9], the Logit-Normal model has the best performance. The application of the HB method in the Logit-Normal model in the Mukomuko District poverty data get estimate that have different value from the direct estimation.
However, some of the two estimators have the same tendency. The model obtained has good flexibility, it can be seen from the plot of the alleged result that is able to follow the pattern of observation data distribution. From these results, there are 3 villages that have a proportion of poor population of more than 40%. The villages are Kota Praja Village, Banjar Sari Village, and Agung Jaya Village. Furthermore, the majority (21 villages) HB estimator have variance less than or equal to the direct estimator variance. This indicates that the estimation using the HB method can improve the estimation of the proportion parameters obtained by using the direct estimation method.

References
[1] Gill J and Torres M 2019 Generalized Linear Models: A Unified Approach (Quantitative Applications in the Social Sciences Book 134) (English: SAGE Publications)
[2] Agresti A 2015 Foundations of Linear and Generalized Linear Models
[3] Fox J and Weisberg S 2019 R Companion to Applied Regression, Third edition. Sage, Thousand Oaks CA Thousand Oaks CA Sage.
[4] Thomson W, Jabbari S, Taylor A E, Arlt W and Smith D J 2019 Simultaneous parameter estimation and variable selection via the logit-normal continuous analogue of the spike-and-slab prior J. R. Soc. Interface
[5] Molenberghs G, Verbeke G and Demétrio C G B 2017 Hierarchical models with normal and conjugate random effects: A review SORT
[6] Tak H and Morris C N 2017 Data-dependent posterior propriety of a Bayesian beta-binomial-logit model Bayesian Anal.
[7] Ghosh J, Li Y and Mitra R 2018 On the use of Cauchy prior distributions for Bayesian logistic regression Bayesian Anal.
[8] Wang C and Blei D M 2018 A general method for robust Bayesian modeling Bayesian Anal.
[9] Wang X, Philip V M, Ananda G, White C C, Malhotra A, Michalski P J, Karuturi K R M, Chintalapudi S R, Acklin C, Sasner M, Bennett D A, De Jager P L, Howell G R and Carter G W 2018 A bayesian framework for generalized linear mixed modeling identifies new candidate loci for late-onset Alzheimer’s disease Genetics
[10] Ropo Ebenezer O and Lougue S 2019 Bayesian generalized linear mixed modeling of breast cancer Iran. J. Public Health
[11] Dietz L R and Chatterjee S 2014 Logit-normal mixed model for Indian monsoon precipitation Nonlinear Process. Geophys.
[12] Franco C and Bell W R 2014 Borrowing Information Over Time in Binomial/Logit Normal Models for Small Area Estimation. Stat. Transit. New Surv. Methodol. Jt. Issue Small Area Estim. 16 563–84
[13] Torabi M and Shokoohi F 2014 Hierarchical Bayes estimation in small area estimation using cross-sectional and time-series data J. Stat. Comput. Simul.
[14] Trevisani M and Torelli N 2017 A Comparison of Hierarchical Bayesian Models for Small Area Estimation of Counts Open J. Stat.
[15] Handayani D, Folmer H, Kurnia A and Notodiputro K A 2018 The spatial empirical Bayes predictor of the small area mean for a lognormal variable of interest and spatially correlated random effects Empir. Econ.
[16] Clement E P 2014 Small Area Estimation with Application to Disease Mapping Int. J. Probab. Stat.
[17] Satopää V A, Baron J, Foster D P, Mellers B A, Tetlock P E and Ungar L H 2014 Combining multiple probability predictions using a simple logit model Int. J. Forecast.
[18] Rao J N K and Molina I 2015 Small Area Estimation: Second Edition
[19] Sunandi E, Agustina D and Fransiska H 2019 Poverty Mapping of the Coastal Areas Using Spatial Empirical Best Linear Unbiased Prediction Method Journal of Physics: Conference Series