Journal of Applied Engineering and Technological Science
Vol 3(2) 2022 : 168-177

CLASSIFICATION OF THE RISK OF COMORBID COVID-19 PATIENT AT BENGKALIS HOSPITAL USING BAYESIAN BINARY LOGISTICS REGRESSION

Muhammad Marizal
Faculty of Science and Technology, Universitas Islam Negeri Sultan Syarif Kasim Riau
m.marizal@uin-suska.ac.id

Received : 24 May 2022, Revised: 28 June 2022, Accepted : 30 June 2022
**Corresponding Author*

ABSTRACT
COVID-19 is an infectious disease caused by SARS-CoV-2 (Severe Acute Respiratory Syndrome Coronavirus 2). This viral infection causes illness with symptoms ranging from mild to severe. The number of deaths from this disease is increasing day by day. A person who is most easily infected with the COVID-19 virus is a person who has a comorbid disease, because the body's immunity decreases due to the impact of a previous illness. The purpose of this study was to determine the comorbid factors that trigger a person's death due to COVID-19. This research uses binary logistic regression with Bayes method parameter estimation. In this study, the predictor variables used were in the form of categories. The results showed that the factors that influence the death of a person on the death of COVID-19 in comorbid diseases are Diabetes Mellitus and Pneumonia.
Keywords : *Binary Logistic Regression, COVID-19, Comorbid*

1. Introduction

Statistics is knowledge related to the collection, presentation, processing, analysis, and various data analysis techniques. Statistics is used as a scientific method to collect, process, summarize, and present research data. In addition, statistics are a way to process data and draw conclusions and careful logistical decisions when processing data (Kurniasih, 2020). Logistic regression analysis is a statistical analysis used to confirm the relationship between variables. Logistic regression can be divided into two, based on the predictor used. That is, simple logistic regression (using one predictor) and multiple logistic regression (using two or more predictors). Meanwhile, logistic regression is divided into two based on the answers used. Namely, binary logistic regression (answers in two categories), multinomial logistics (answers in three or more categories), and ordinal logistics (answers in three or more categories but layered) (Albana, 2013).

Binary logistic regression analysis is a logistic receipt reg analysis whose answer is binary. This binary logistic regression model is used when the answer variable creates two categories with values Y = 1 (success) and Y = 0 (failure) according to the Bernoulli distribution for each single observation (Nugraha, 2013). Bayesian binary logistic regression analysis is a binary logistic regression analysis that uses the Bayesian method (Apsari et al., 2013). Bayesian method is a parameter estimation method that assumes that the parameters are random variables with a certain distribution. The Bayesian method combines the likelihood function and the prior distribution of the previous parameters to obtain a posterior distribution. The posterior distribution is the basis for estimating the parameters (Shobri et al., 2021).

Coronavirus Disease 2019 (COVID-19) is an infectious disease caused by a new type of corona virus that was discovered in 2019. Infection with this virus can cause illness with symptoms ranging from mild to severe (Yuliana, 2020). COVID-19 can cause respiratory infections, pneumonia, and other deaths. According to KPC PEN (Committee for the Management of COVID-19 and National Economic Recovery), in July 2021 there were 2,284,084 people infected with COVID-19 in Indonesia. Of the total, 1,928,274 cases have been declared cured and the death toll has reached 60,582 people. Considering the higher mortality rate of COVID-19 comorbid patients in Indonesia, a study was conducted on what comorbid factors could be the cause of the risk of death in COVID-19 comorbid patients. The response variable in this study

was the patient's discharge status, with two options, namely the patient was declared cured or died. This variable follows the Bernoulli distribution, so the binary regression method can be used.

Logistic regression model was used recently in different research. Logistic regression model and artificial neural network were considered with several independent variables to predict mortality due to COVID-19 pandemic in 32 European countries (Elhag et al., 2021). Logistic regression analysis was utilized to identify COVID-19 lockdown measures related risk factors in various studies (Rossi et al., 2020; Alfawaz et al., 2021; ting et al., 2021; Jacob et al., 2021). Binary logistic regression was used to investigate customer impression towards online delivery (Mehrolia et al., 2021). Multiple logistic regression model was used to evaluate the association between anxiety and various factors during COVID-19 (Fu et al., 2021; Jiang et al., 2020); Liu et al., 2021). Models' coefficients were estimated by utilizing different methods such as maximum likelihood method, Bayesian method and least square method (Sindhu et al., 2021; Pandey et al., 2021; Zou et al., 2021; Zou et al., 2021; Jairi et al., 2021).

Research on COVID-19 mortality factors is a hot topic that continues to be discussed today, such as research conducted by Research (Shobri et al., 2021) showing that COVID-19 mortality factors can use the Bayesian binary logistic regression method. When solving problems in Bayesian methods, one must first determine the prior and posterior probabilities. The results of this study show that the mortality rate of comorbid patients due to COVID-19 is very significant. So the more comorbidities a person has, the easier it is for that person to die from COVID-19.

Research (Satria et al., 2020) which shows that the factors of death from COVID-19 are men, old age, diabetes and hypertension (comorbid factors). As research (Lestari et al., 2020) which results show diabetes mellitus is one of the comorbid factors among many comorbid factors that increase the risk of severity and death in COVID-19 patients. Research that uses Bayesian binary logistic regression to determine the influencing factors is carried out by research (Nadhifah & Yasin, 2012) (Sudirohusodo et al., 2020). Research (Nadhifah & Yasin, 2012) shows that from the results of the binary logistic regression model, the factors that affect a baby born with normal conditions are age and maternal hemoglobin levels. Furthermore, research conducted by (Sudirohusodo et al., 2020) showed that the best models for the factors that influence the type of breast cancer malignancy are age, location of breast cancer, breast cancer chemotherapy, percentage of breast cancer patients metastases and marital status.

Research on Bayesian binary logistic regression has been carried out by (Nadhifah & Yasin, 2012; Sudirohusodo et al., 2020; Susilo et al., 2020; Widayani et al., 2020). Where it is explained that the Bayes method is a method used to estimate the parameters of a regression model that uses probability interpretation by combining sample data with other information provided previously (prior). When performing the convergence test, use the Bayes method for model parameters whose values have been estimated, by observing the results displayed by the trace plot. Next, the form of the probability density function will be shown for each estimated model parameter obtained from the Gibbs sampler process (Susilo et al., 2020). Research that uses $press's\ Q$ as hypothesis testing is (Shobri et al., 2021; Susilo et al., 2020). Which shows that the value of $press's\ Q$ is greater than the value of $\chi^2_{0,05;1}$ (219,88 > 3,841) then the model is accurate and stable (Shobri et al., 2021).

According to a report from the National Disaster Management Agency (BNPB) July 4, 2021, Riau Province ranks eighth with 463 daily cases of COVID-19. In Bengkalis Regency, one of the governments in Riau, where COVID-19 has been confirmed since last July 20, the number of cases has exceeded 83 cases. Head of the Bengkalis District Health Office, Ersan Saputra Th, said the additional 83 people came from 10 sub-districts, the highest being in 2 sub-districts, namely Bengkalis Regency with 24 people and Mandau 22 people.

The problems that arise in recognizing the spread of COVID-19 are various factors that are considered the cause of death. Therefore, a suitable method is needed to determine the trigger. The causes (risks) used in this study are comorbid factors, namely Hypertension, Diabetes Mellitus, COPD, Pneumonia. Based on the description above, the author is interested in conducting a study with the title Risk Classification of Comorbid COVID-19 Patients at Bengkalis Hospital Using Bayesian Binary Logistics Regression.

2. Literature Review

Coronavirus Diases 2019 (Covid-19) is a disease caused by the SAR-CoV-2 virus (Severe Acute Respiratory Syndrome Coronavirus 2) where this virus has a positive single strain RNA, encapsulated and not segmented. Infection from Covid-19 can be in the form of mild, moderate or severe symptoms. The main clinical symptoms that appear are fever (temperature over 38°C), cough and difficulty breathing. In addition, it can be accompanied by severe shortness of breath, fatigue, myalgia, gastrointestinal symptoms such as diarrhea and other respiratory tract symptoms. In severe cases this Covid-19 infection can cause critical illness and death (Yuliana, 2020). The death of Covid-19 patients is not only caused by Covid-19, there are several other supporting factors that can cause the risk of death of Covid-19 patients to increase. Based on research conducted by (Susilo et al., 2020) and (Hidayani et al., 2020) there are several factors that influence the risk of death of Covid-19 patients, including:

1. Hypertension

Hypertension or high blood pressure is a condition when blood pressure reaches 130/80 mmHg or more. Hypertension which currently records the highest mortality rate in comorbid COVID-19 patients is hypertension when the blood pressure is above 140/90 mmHg (Research & Siagian, 2020).

2. Diabetes Mellitus

Diabetes mellitus is a chronic disease characterized by high blood sugar, or usually this disease is called sugar disease or diabetes. Many comorbid factors in COVID-19 patients occur due to weak insulin production so that the body's ability to respond to insulin is reduced (Lestari et al., 2020).

3. COPD

Chronic obstructive pulmonary disease (COPD) is inflammation of the lungs that develops over a long period of time. COPD is characterized by difficulty breathing, coughing up phlegm and wheezing. COPD is an unhealthy behavior so it is an estimate of one of the comorbidities in the death of COVID-19 patients (Satria et al., 2020).

4. Pneumonia

Pneumonia (wet lungs) is an inflammation of the lungs caused by infection. Pneumonia can cause mild to severe symptoms. Some of the symptoms commonly experienced by pneumonia sufferers are cough with phlegm, fever, and shortness of breath (Klein, 2020).

3. Research Methods

The dependent variable used in this study is the patient's discharge status, where there are two possibilities, namely patients who are declared cured or died at Bengkalis Hospital from January to September 2021. The independent variables used are Hypertension, Diabetes Mellitus, COPD and Pneumonia. The method used in this research is Multiple Linear Logistic Regression and Bayesian Binary Logistic Regression.

Multiple linear regression is an equation model used to explain the relationship between the dependent variable (Y) and two or more independent variables (X). The purpose of multiple linear regression is to predict whether the value of the independent variable is known and can determine the relationship between the dependent variable and the independent variable (Teti et al., 2015). Estimating the regression parameters using the Ordinary Least Square (OLS) method (Harlianingtyas & Iriawan, 2015).

In testing the regression parameters, there are two tests that need to be done, namely the simultaneous test and the partial test. Testing the regression model parameters simultaneously using the F distribution approach. Simultaneous significance test is to test all independent variables in the overall model simultaneously. This test was conducted to see whether the independent variables as a whole had a significant effect on the dependent variable. In this case it is necessary to compare the values of and (Sartika & Debataraja, 2020). Partial testing using the t distribution approach. Spatial regression coefficient test is used to prove the effect of independent variables on the dependent variable. The t-test was carried out to partially determine the significance of each estimator parameter assuming the other independent variables were constant (Utami et al., 2017).

Logistic regression analysis is a statistical analysis used to see the relationship between variables. Based on the predictor variables used, logistic regression is divided into two, namely simple logistic regression (with one predictor variable) and multiple logistic regression (with two or more predictor variables). Meanwhile, based on the response variables used, logistic regression is divided into two, namely binary logistic regression (response variable with two categories) and multinomial and ordinal logistic regression (response variable with more than two categories) (Albana, 2013).

Binary logistic regression analysis is a logistic regression analysis with a binary or dichotomous response variable. The binary logistic regression model is used if the response variable produces two categories with values Y=1 (success) and Y=0 (failure), so it follows the Bernoulli distribution for each single observation (Nugraha, 2013).

In general, the logistic regression model is as follows:

$$\pi(x) = \frac{\exp(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k)}{1 + \exp(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k)}$$

When $0 < \pi(x) < 1$.

Markov Chain Monte Carlo (MCMC)

Markov Chain Monte Carlo (MCMC) is a method used to determine parameter values of a difficult analytical integration. The MCMC method used is the Gibbs Sampler. In this study using the Gibbs Sampler method, because this study uses the Bayesian method. The Gibbs Sampler method applied in this study uses the WinBugs software (Robert & Ntzoufras, 2012).

The Gibbs Sampler Algorithm is as follows:

1. Determine the parameters to be used
2. Modeling the Gibbs Sampler algorithm

$$g(x) = ln\frac{\pi(x)}{[1-\pi(x)]} = (\beta_0 + \beta_1 x_1 + \cdots + \beta_k x_k)$$

3. Enter the data according to the data obtained from the hospital.
4. Specifies the prefix value of each parameter.
5. For samples in posterior analysis use $\{\beta^1, \beta^2, \cdots, \beta^L\}$.
6. Doing Iteration Update
7. Determine the mean and variance of the posterior distribution

In the Bayes method, there are several ways to see the convergence of model parameters, including Trace plots, Density plots or by comparing the Monte Carlo Error (MC Error) value with the standard deviation of the model parameters. The model parameter values are said to be convergent if the plot distribution is stable spread between two parallel or stationary horizontal lines and the resulting Density plot image resembles a normal distribution curve (Satria et al., 2020).

Parameter significance test was conducted to determine the feasibility of the model by testing the significance of each independent variable used in the model. The absence of treatment effect is stated by the null hypothesis whose parameter value causes the treatment to be zero. The opposite hypothesis which states that a treatment is not worth zero is called the alternative hypothesis. Hypothesis testing for Bayesian uses a credible interval whose simplest form can be addressed by the quantiles of 2.5% and 97.5%. If the mean value contains a value of zero, it can be concluded that there is no effect of treatment on the response variable (Robert & Ntzoufras, 2012)

4. Results

4.1 Multiple Linear Regression Model

Prior to modeling using the Bayes method, multiple regression will be carried out as a model using OLS estimation. The results of the multiple regression estimation using SPSS software can be seen in the following table:

Table 1 - Parameter Estimation of Multiple Linear Regression Model

Variable	Estimate
Intercept	-0,047
Hypertension (X1)	0,179
DM (X2)	0,248
PPOK (X3)	0,259
Pneumonia (X4)	0,479

Based on the estimation of the regression model parameters, the multiple regression model can be written as follows:

$$Y = -0{,}047 + 0{,}248X_2 + 0{,}479X_4$$

Multiple linear regression models can be formed by testing significant parameters simultaneously and partially.

Table 2 - Simultaneous Test

Variance	df	Number of square	Sum Square	F
Regression	4	7,079	1,770	9,381
Error	217	40,749	0,189	
Total	221	47,828		

The linear regression model was used to determine which independent variables had a significant effect on the mortality of COVID-19 patients. Based on the results of the *F* test, the value of $F_{Count} = 9{,}381$ and $F_{table} = 2{,}41$ so that it can be concluded that simultaneously there is at least one influence between the independent variable and the dependent variable.

In addition to the F test on the linear regression model, a partial test was also carried out using the t test. Based on the test results, it is obtained that all the regression coefficients of each independent variable and the intercept are significant to the model because all values of $t\_count > t\ table$. Based on this, it can be concluded that the independent variables that affect the mortality of COVID-19 patients are Diabetes Mellitus (DM) and Pneumonia.

Table 3 - Partial Test

Variable	t_{Count}	Conclusion
Hypertension (X1)	1,770	no Sig.
DM (X2)	2,808	Significant
PPOK (X3)	1,359	No Sig.
Pneumonia (X4)	5,791	Significant

The decision is that the hypothesis rejects H_0 meaning that the influence of the independent variable affects the mortality of COVID-19 patients with comorbid diseases

4.2 Binary Regression Model with Bayes Method

Parameter estimation is done using Bayes method with the help of WinBugs software. Elimination of the regression parameter hypothesis was carried out using a 95% confidence interval approach for each parameter. The 95% confidence interval was calculated with the lower limit being the 2.5% quantile and the upper limit being the 97.5% quantile. The parameter is declared significant if the 95% confidence interval on the parameter does not contain a zero value. Significant parameters indicate that the independent variable has an effect on the response and insignificant parameters indicate that the independent variable has no effect on the response. The following is the estimated value and the conclusion.

Table 4 - Parameter Estimation Value

Variable	Parameters	Mean	2,5% Quantil	97,5% Quantil	Sig.
Intercept	β_0	-2,677	-3,982	-1,332	-
Hypertension (X1)	β_1	0,9116	-0,5054	2,134	No
DM (X2)	β_2	1,187	0,1441	2,396	Yes
PPOK (X3)	β_3	0,4092	-3,673	2,808	No
Pneumonia (X4)	β_4	2,381	1,187	3,614	Yes

Table 4 shows that the influential variables are x_2 namely Diabetes Mellitus and x_4 namely Pneumonia. While the variable x_1 is Hypertension x_3 namely COPD (Chronic Obstructive Pulmonary Disease) has no effect on the response variable. Therefore, these two variables were not included in the binary logistic regression model. So the model formed is:

$$\pi(x) = \frac{\exp\left(-2{,}677 + 1{,}187X_2 + 2{,}381X_4\right)}{1 + \exp\left(-2{,}677 + 1{,}187X_2 + 2{,}381X_4\right)}$$

4.3 Model Parameter Convergence Test

After the estimated model parameters are obtained, then the convergence test is carried out by observing the results of the trace plot and density plot. The following are the results of the trace plot and density plot using the WinBUGS software application

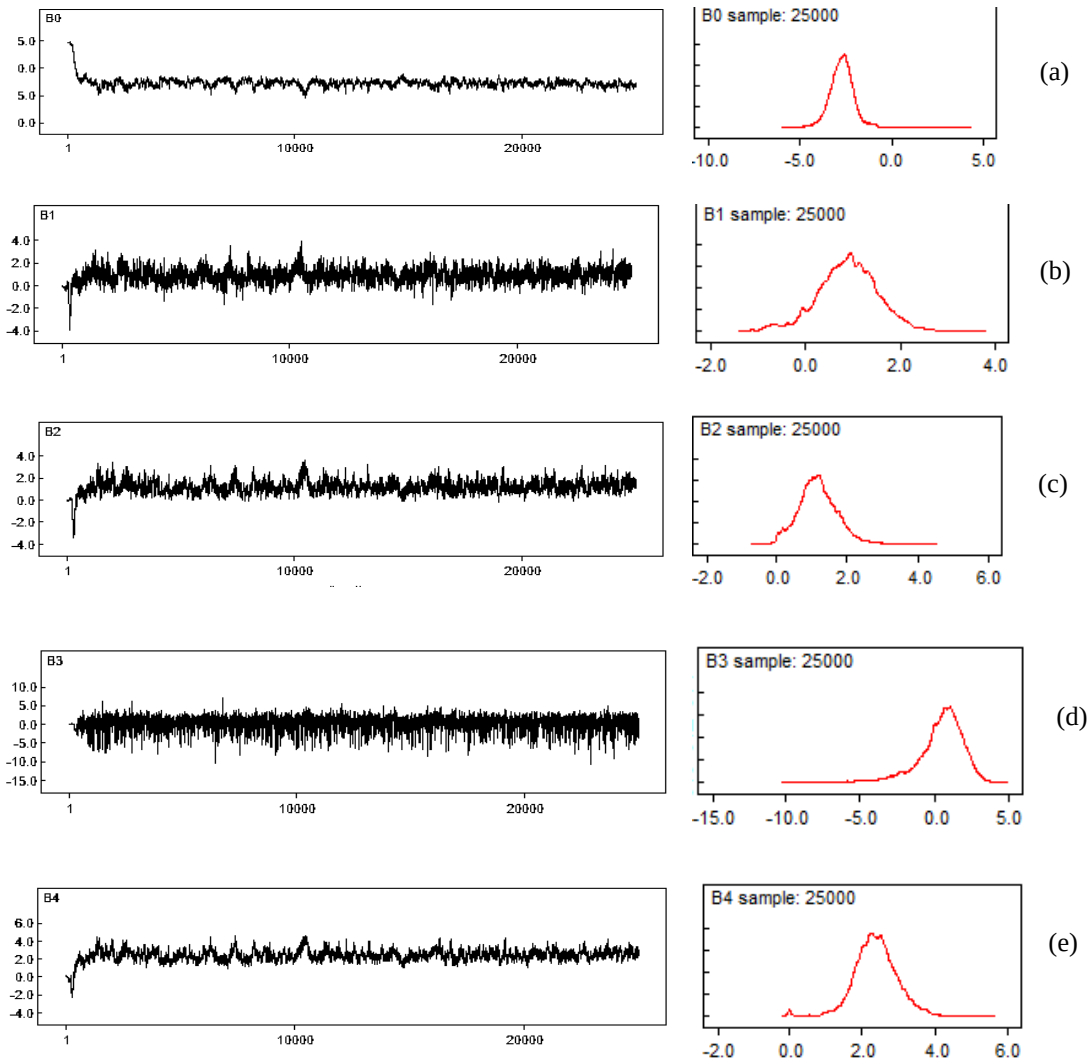

Fig. 1. Results of History Trace Plot and Density Plot for Parameter (a). β_0 (b) β_1 (c).β_2 (d).β_3 (e).β_4

Based on Figure 1 above, it can be concluded that the convergence assumption is fulfilled because the iteration data distribution is stable and lies between two parallel horizontal lines and the resulting density plot is quite good because it has a pattern that tends to be smooth in the form of a normal distribution curve.

4.4 Model Interpretation

To find out how much the predictor variable has an effect on the response variable is to look at the odds ratio value of the model parameters. The value of the odds ratio of the model parameters can be seen in the following table.

Table 5 - Odds Ratio Value

Variable	β	$exp\ (\beta)$
DM (X2)	0,9127	3,898
Pneumonia (X4)	0,9312	3,221

Based on Table 5 the interpretation of the odds ratio of each variable is the diabetes mellitus variable, the odds ratio value is 3.898. From this value, it can be interpreted that comorbid patients suffering from diabetes mellitus have a 3.898 times greater risk of dying from COVID-19 compared to comorbid patients who do not suffer from diabetes mellitus. As for the pneumonia variable, the odds ratio value was 3.221. From this value, it can be interpreted that comorbid patients who suffer from pneumonia have a risk of dying from COVID-19 which is 3,221 times greater than that of comorbid patients who do not suffer from pneumonia.

4.5 Classification Accuracy Analysis

The logistic regression model is said to be good and stable where the classification accuracy value is close to 1 and the value of press^\'s Q is greater than the value of the chi-square table. The value of accuracy and press^\'s Q is obtained by looking at the accuracy of the model in classifying or grouping. If the probability value is greater than or equal to 0.5 then the patient's discharge status is grouped into the dead category, whereas if the logit value is less than 0.5 then the patient's discharge status is grouped into the cured category. There were 64 errors in the estimation of classification, namely 56 and 8 which were grouped in the wrong category from a total of 221 respondents.

Based on the value of classification accuracy and *press's Q* above, the model can classify the data correctly by 71.04% and the value of *press's Q* is greater than the value of $\chi^2_{0\,05;1}$table ($\chi^2_{0\,05;1} = 3{,}841$) which is $39{,}135 > 3{,}841$. This means that the model is accurate and statistically stable in classifying the risk of death for Covid-19 comorbid patients.

5. Discussions

In this study, only two independent variables affected the mortality variable of COVID-19 patients. The independent variables that have an effect on this study are Diabetes Mellitus (X_2) and Pneumonia (X_4), while the variables that have no effect on the mortality of comorbid COVID-19 patients are Hypertension (X_1) and COPD (X_3). The dependent variable used in this study is the Patient's Exit Status (Y). Several previous studies have conducted research on Binary Logistics Regression with Bayes Method. Research on COVID-19 mortality factors is a hot topic that continues to be discussed today, such as research conducted by [7] which showed that COVID-19 mortality factors were men, old age, diabetes and hypertension (comorbid factors). As research [8] whose results show diabetes mellitus is one of the comorbid factors among many comorbid factors that increase the risk of severity and death in COVID-19 patients. Research that uses Bayesian binary logistic regression to determine the influencing factors is carried out by research (Nadhifah & Yasin, 2012; Sudirohusodo et al, 2020).

Research (Nadhifah & Yasin, 2012) shows that from the results of the binary logistic regression model, the factors that affect a baby born with normal conditions are age and maternal hemoglobin levels. Furthermore, research conducted by Sudirohusodo et al., (2020) showed that the best models for the factors that influence the type of breast cancer malignancy are age, location

of breast cancer, breast cancer chemotherapy, percentage of metastases in breast cancer patients and marital status.

The research of Sobri et al., (2021) shows that the factors of COVID-19 mortality can use the Bayesian method of binary logistic regression. When solving problems in Bayesian methods, one must first determine the prior and posterior probabilities. The results of this study show that the mortality rate of comorbid patients due to COVID-19 is very significant. So the more comorbidities a person has, the easier it is for that person to die from COVID-19.

Research on Bayesian binary logistic regression has been carried out by (Nadhifah & Yasin, 2012; Sudirohusodo et al., 2020; Susilo et al., 2020; Widayani et al., 2020). Where it is explained that the Bayes method is a method used to estimate the parameters of a regression model that uses probability interpretation by combining sample data with other information provided previously (prior). When performing the convergence test, use the Bayes method for model parameters whose values have been estimated, by observing the results displayed by the trace plot. Next, the form of the probability density function will be shown for each estimated model parameter obtained from the Gibbs sampler process (Susilo et al., 2020). Research that uses press^'s Q as hypothesis testing is (Shobri et al., 2021; Susilo et al., 2020). Which shows that the value of press^' s Q is greater than the value of _(0,05;1)^2 (219.88 > 3,841) then the model is accurate and stable (Shobri et al., 2021). Based on the research above, there are the same independent variables in influencing the mortality of COVID-19 patients, namely Hypertension and Diabetes Mellitus and there are also different variables. This is due to the location/area of the study and the different years.

6. Conclusion

Based on the discussion that has been described, it can be concluded that comorbid Diabetes Mellitus (x_2) and Pneumonia (x_4) have a significant influence on the risk of death of Covid-19 patients. The logistic regression model obtained for the risk of death of Covid-19 patients is:
Multiple linear regression model with least squares method

$$Y = -0{,}047 + 0{,}248X_2 + 0{,}479X_4$$

Binary logistic regression model with Bayes method

$$\pi(x) = \frac{\exp\left(-2{,}677 + 1{,}187X_2 + 2{,}381X_4\right)}{1 + \exp\left(-2{,}677 + 1{,}187X_2 + 2{,}381X_4\right)}$$

with a classification accuracy model of 71.04%. Thus, it can be said that the model is feasible in classifying the risk of death for COVID-19 comorbid patients who are declared cured of COVID-19 comorbid patients.

References

Albana, M. (2013). Aplikasi Regresi Logistik Ordinal Untuk Menganalisa Tingkat Kepuasan Pengguna Jasa Terhadap Pelayanan Di Statiun Jakarta Kota. *Matematika, Program Studi Matematika, Fakultas Pengetahuan*, *1*, 56–72.

Apsari, W., Yasin, H., & Sugito. (2013). Estimasi Parameter Regresi Logistik Multinomial Dengan Metode Bayes. *None*, *2*(1), 79–88.

Harlianingtyas, I., & Iriawan, N. (2015). *Pemodelan Resiko Pembiayaan Griya Bank Mandiri Syariah Dengan Metode Bayesian Regresi Logistik* (pp. 606–613).

Hidayani, W. R., Studi, P., & Masyarakat, K. (2020). *Faktor Faktor Risiko Yang Berhubungan Dengan COVID 19 : Literature Review*. *4*(2), 120–134.

Klein, F. (2020). Risikofaktor Komorbiditäten bei COVID-19- Erkrankung. In *Pneumologie* (Vol. 74, Issue 10, p. 640). https://doi.org/10.1183/13993003.00547-2020

Kurniasih, N. (2020). Peran Statistik Dalam Penelitian. *Al-Munqidz : Jurnal Kajian Keislaman*, *8*(2), 273–279. https://doi.org/10.52802/amk.v8i2.254

Lestari, N., Ichsan, B., Kedokteran, F., Surakarta, U. M., Lestari, K. N., & Alamat, M. P. H. (2020). *Diabetes Melitus Sebagai Faktor Risiko Keparahan Dan Kematian Pasien Covid-19 : Meta-Analisis Diabetes Mellitus As A Risk Factor For Severity And Mortality Of*

Covid-19 : A Meta-Analysis. *13*(1), 83–94. https://doi.org/10.23917/biomedika.v13i1.13544

Nadhifah, L., & Yasin, H. (2012). *Analisis Faktor-Faktor Yang Mempengaruhi Bayi Berat Menggunakan Metode Bayes (Studi Kasus di Rumah Sakit Umum Daerah Kota Semarang)*. *1*, 125–134.

Nugraha, J. D. (2013). *Pengantar Analisis Data Kategorik* (Issue 0274). https://dspace.uii.ac.id/bitstream/handle/123456789/23388/Pengantar Analisis Data Kategorik Deepublish -free view.pdf?sequence=4&isAllowed=y

Penelitian, A., & Siagian, T. H. (2020). *CORONA DENGAN DISCOURSE NETWORK ANALYSIS*. *09*(02), 98–106.

Robert, C., & Ntzoufras, I. (2012). Bayesian Modeling Using WinBUGS. *CHANCE*, *25*(2), 60–61. https://doi.org/10.1080/09332480.2012.685377

Sartika, I., & Debataraja, N. N. (2020). Analisis Regresi Dengan Metode Least Absolute Shrinkage And Selection Operator (LASSO) Dalam Mengatasi Multikolinieritas. *Buletin Ilmiah Math Stat Dan Terapan (Bimaster)*, *09*(1), 31–38.

Satria, R. M. A., Tutupoho, R. V., & Chalidyanto, D. (2020). Analisis Faktor Risiko Kematian dengan Penyakit Komorbid Covid-19. *Jurnal Keperawatan Silampari*, *4*(1), 48–55. https://doi.org/10.31539/jks.v4i1.1587

Shobri, M. Q., Yanuar, F., & Devianto, D. (2021). Covid-19 Patient Mortality Risk Classification Using Bayesian Binary Logistic Regression. *Jurnal Matematika, Statistika Dan Komputasi*, *18*(1), 150–160. https://doi.org/10.20956/j.v18i1.14268

Sudirohusodo, W., Meliza, B. D., & Bustan, M. N. (2020). *Analisis Regresi Logistik dengan Metode Penduga Bayes untuk Menentukan Faktor-Faktor Yang Berpengaruh Terhadap Kejadian Kanker Payudara (Studi Kasus : Pasien Kanker Payudara di Rumah sakit*. *2*(2), 52–61. https://doi.org/10.35580/variansiunm12900

Susilo, A., Rumende, C. M., Pitoyo, C. W., Santoso, W. D., Yulianti, M., Sinto, R., Singh, G., Nainggolan, L., Nelwan, E. J., Khie, L., Widhani, A., Wijaya, E., Wicaksana, B., Maksum, M., Annisa, F., Jasirwan, O. M., Yunihastuti, E., Penanganan, T., New, I., … Cipto, R. (2020). *Coronavirus Disease 2019 : Tinjauan Literatur Terkini Coronavirus Disease 2019 : Review of Current Literatures*. *7*(1), 45–67.

Teti, M. S., Yanuar, F., & Yozza, H. (2015). Analisis Regresi Logistik Dengan Metode Penduga Bayes Untuk Menentukan Faktor-Faktor Yang Mempengaruhi Kejadian Bayi Berat Badan Lahir Rendah. *Jurnal Matematika UNAND*, *4*(1), 53. https://doi.org/10.25077/jmu.4.1.53-60.2015

Utami, T. W., Rohman, A., & Prahutama, A. (2017). Pemodelan Regresi Berganda Dan Geographically Weighted Regression Pada Tingkat Pengangguran Terbuka Di Jawa Tengah. *Media Statistika*, *9*(2), 133. https://doi.org/10.14710/medstat.9.2.133-147

Yuliana, Y. (2020). Corona virus diseases (Covid-19): Sebuah tinjauan literatur. *Wellness And Healthy Magazine*, *2*(1), 187–192. https://doi.org/10.30604/well.95212020

Sindhu, T. N., Shafiq, A., Al-Mdallal, Q.M. (2021). Exponentiated transformation of Gumbel Type-II distribution for modeling COVID-19 data, *Alexandria Eng. J. 60* (1), 671–689.

Elhag, A.A., Aloafi, T.A., Jawa, T.M., Sayed-Ahmed, N., Bayones, F.S., Bouslimi, J. (2021). Artificial neural networks and statistical models for optimization studying COVID-19, *Results Phys. 25*, 104274, https://doi.org/10.1016/j.rinp.2021.104274.

Rossi, R., Socci, V., Talevi, D., Mensi, S., Niolu, C., Pacitti, F., Di Marco, A., Rossi, A., Siracusano, A., Di Lorenzo, G. (2020). COVID-19 pandemic and lockdown measures impact on mental health among the general population in Italy, Front. *Psychiatry 11*, https://doi.org/10.3389/fpsyt.2020.00790.

Alfawaz, H., Yakout, S.M., Wani, K., Aljumah, G.A., Ansari, M.G.A., Khattak, M.N.K., Hussain, S.D., Al-Daghri, N.M.(2021). Dietary intake and mental health among Saudi adults during COVID-19 lockdown, Int. *J. Environ. Res. Public Health 18* (4), 1653, https://doi.org/10.3390/ijerph18041653.

Ting, D.S.J., Krause, S., Said, D.G., Dua, H.S. (2021). Psychosocial impact of COVID-19 pandemic lockdown on people living with eye diseases in the UK, *Eye 35* (7), 2064–2066.

Jacob, L., Smith, L., Armstrong, N. C., Yakkundi, A., Barnett, Y., Butler, L., ... & Tully, M. A. (2021). Alcohol use and mental health during COVID-19 lockdown: A cross-sectional study in a sample of UK adults. *Drug and alcohol dependence*, 219, 108488

Mehrolia, S., Alagarsamy, S., V.M. Solaikutty, V. M. (2021). Customers response to online food delivery services during COVID-19 outbreak using binary logistic regression, Int. *J. Consumer Studies 45* (3), 396–408.

W. Fu, W., Yan, S., Zong, Q., Anderson-Luxford, D., Song, X., Lv, Z., Lv, C. (2021). Mental health of college students during the COVID-19 epidemic in China, *J. Affect. Disord. 280* 7–10.

Jiang, W., Liu, X., Zhang, J., Feng, Z. (2020). Mental health status of Chinese residents during the COVID-19 epidemic, *BMC Psychiatry 20* (1), 1–14.

Liu, Z., Liu, R., Zhang, Y., Zhang, R., Liang, L., Wang, Y., Wei, Y., Zhu, R., Wang, F. (2021). Latent class analysis of depression and anxiety among medical students during COVID-19 epidemic, *BMC Psychiatry 21* (1), https://doi.org/10.1186/s12888-021-03459-w.

Pandey, A. K., Singh, G, N., Sayed-Ahmed, N., Abu-Zinadah, H. (2021). Improved estimators for mean estimation in presence of missing information, *Alexandria Eng. J. 60* (6), 5977–5990.

Zuo, L., Zhang, H., HaiYing Wang, Sun, L. (2021). Optimal subsample selection for massive logistic regression with distributed data, *Comput. Statistics 36* (4), 2535–2562.

Jairi, I., Fang, Y., & Pirhadi, N. (2021). Application of Logistic Regression Based on Maximum Likelihood Estimation to Predict Seismic Soil Liquefaction Occurrence. *Human-Centrice Intellegent Systems.*