Assessing the Effect of Different Covariates Distributions on Parameter Estimates for Multinomial Logistic Regression (MLR)

Hamzah Abdul Hamid¹,a), Siti Raudhah Ismail²,b), Sahimel Azwal Sulaiman³,c) and Nor Azrita Mohd Amin¹,d)

¹Institute of Engineering Mathematics, Universiti Malaysia Perlis, Pauh Putra Main Campus, 02600 Arau, Perlis, Malaysia
²Kolej Matrikulasi Kedah, Kementerian Pendidikan Malaysia, 06010 Changlun, Kedah Darul Aman, Malaysia
³Faculty of Muamalah & Management, Kolej Universiti Islam Perlis, Taman Seberang Jaya, Fasa 3, Seberang Ramai, 02000 Kuala Perlis, Perlis, Malaysia

Corresponding author: hamzahhamid@unimap.edu.my

Abstract. In fitting a multinomial logistic regression model, one of the most important part is estimating the parameter. In Multinomial Logistic Regression (MLR), Maximum Likelihood Estimation (MLE) method is used to estimate the parameters. MLE is the suitable method to be applied to the problems associated with categorical response variables since it has several benefits such as sufficiency, consistency, efficiency and parameterization invariance. This study investigates the different type of continuous distributions (normal, negatively skewed, positively skewed) on parameter estimation via Monte Carlo simulation. From the simulation result, it shows that as the sample size increases, the effect of covariate distribution reduces. The estimated parameter also less affected for model with normal covariate distribution. At sample size 300 and above, the estimated parameter with normal covariate distribution is considered as close to the true parameter value. Interestingly, for the positively skewed, the estimated parameter also obtained unbiased parameter at sample size 300 and above. However, for negatively skewed, it requires a larger sample size to get closer to the true parameter value. The estimated parameters deviate too far from the true parameter at small sample size. As expected, as sample size increases the parameter estimates for all distributions are getting close to the true parameter value. Lastly, the distribution for MLR with more than one covariate give the same effect as the MLR model with only one covariate on parameter estimations.

1. Introduction

The logistic regression is one of the GLMs and it is used to model the relationships between a categorical response variables, Y and one or more independent variables, X=(x₁,x₂,…..xₚ). The response variables Y may have two or more categories and these categories can be nominal or ordinal. If Y has only two categories, it is called binary logistic regression. However, when Y has more than two nominal categories, the multinomial logistic regression (MLR) is more relevant to be applied, while in the situation when Y has more than two categories and the categories can be ordered, the ordinal logistic regression is more appropriate to be applied. The independent variables in logistic regression

[1]
model can be quantitative, qualitative or both (mixed). In logistic regression, the maximum likelihood method is usually applied to estimate the parameter.

The distribution of the covariates may affect the estimation of parameter of statistical model. [1] reported that ANOVA and Kruskal-Wallis tests are affected when the data is not normal. They reported that both tests are affected by the kurtosis of the error distribution, but less affected by the skewness. The simulation results of [2] showed that the presence of kurtosis and small sample size are affected the t-test for simple linear regression model with non-normal errors.

The investigation by [3] shows that the parameter estimates for MLR model with a single covariate affected when data is not normally distributed. The parameter estimates are deviated too far from the true parameter when the distribution of covariate is negatively skewed and required very large sample size to get close to the true parameter value. The parameter estimates for MLR with categorical covariate getting closer to the true parameter when the sample size is 300 and above. However, the effect of distribution on MLR model with more than one covariate and minimum sample size required to minimize the effect of covariate distribution is still vague. Therefore, the objective of this simulation study is to investigate the minimum sample size required to minimize the effect of covariate distributions for MLR model with more than one covariates on parameter estimation.

2. Methodology

2.1. Multinomial Logistic Regression

In multinomial logistic regression, let \( Y \) be an outcome variable with \( c \) possible value \( (0,1,\ldots,c-1) \) with \( Y=0 \) be the reference category. Let \( X=(x_1,x_2,\ldots,x_p) \) be the independent variables. Thus, the conditional probabilities of each outcome category can be written as [4]:

\[
P(Y = 0 \mid x) = \frac{1}{1 + e^{g_0(x) + g_1(x) + \cdots + g_{c-1}(x)}}
\]

\[
P(Y = 1 \mid x) = \frac{e^{g_1(x)}}{1 + e^{g_0(x) + g_1(x) + \cdots + g_{c-1}(x)}}
\]

\[
P(Y = c - 1 \mid x) = \frac{e^{g_{c-1}(x)}}{1 + e^{g_0(x) + g_1(x) + \cdots + g_{c-1}(x)}}
\]

It follows that the logit function of category \( j \):

\[
g_j(x) = \ln \left( \frac{P(Y = j \mid x)}{P(Y = 0 \mid x)} \right) = \beta_{j0} + \beta_{j1}x_1 + \ldots + \beta_{jp}x_p \quad \text{for} \; j=1,2,\ldots,c-1
\]

2.2. Simulation Procedures

The simulation study is used to evaluate the effect of distribution for MLR with more than one covariate and the sample size required in order to estimate the parameters. The Monte Carlo simulation is used and the sample sizes of 50, 100, 150, 300, 500, 1000, 1500, 3000 and 5000 are considered in this investigation. The simulation procedure was carried out using R, an open source programming software (version 3.5.1). The data will be generated using the same technique as in [5]. Table 1 shows the distribution of covariates and true parameter values for the multinomial logistic regression model. At least one of the covariate distributions in this model is N(0,1). This is to avoid
the imbalanced data set generated for the dependent variables. The coefficient values are chosen to represent an odds ratio for each covariate is greater and lower than 1. Odds ratio greater than 1 indicates exposure associated with higher odds of the outcome while the odds ratio lower than 1 indicates exposure associated with lower odds of outcome.

### Table 1. The distribution of covariates and true parameter values

| Model | Covariate Distribution | Skewness      | Kurtosis     | Coefficient                  |
|-------|------------------------|---------------|--------------|-----------------------------|
| A     | N(0,1) and N(0,1)      | 0.000 and 0.000 | 2.996 and 2.996 | $\beta_{10} = -2.10$        |
|       |                        |               |              | $\beta_{11} = -0.35$        |
|       |                        |               |              | $\beta_{12} = 1.08$         |
|       |                        |               |              | $\beta_{20} = -1.90$        |
|       |                        |               |              | $\beta_{21} = -0.21$        |
|       |                        |               |              | $\beta_{22} = 1.69$         |
| B     | N(0,1) and Beta (12,1) | 0.000 and -1.577 | 2.996 and 6.108 | $\beta_{10} = -2.10$        |
|       |                        |               |              | $\beta_{11} = -0.35$        |
|       |                        |               |              | $\beta_{12} = 1.08$         |
|       |                        |               |              | $\beta_{20} = -1.90$        |
|       |                        |               |              | $\beta_{21} = -0.21$        |
|       |                        |               |              | $\beta_{22} = 1.69$         |
| C     | N(0,1) and $\chi^2(4)$ | 0.000 and 1.405 | 2.996 and 5.931 | $\beta_{10} = -2.10$        |
|       |                        |               |              | $\beta_{11} = -0.35$        |
|       |                        |               |              | $\beta_{12} = 1.08$         |
|       |                        |               |              | $\beta_{20} = -1.90$        |
|       |                        |               |              | $\beta_{21} = -0.21$        |
|       |                        |               |              | $\beta_{22} = 1.69$         |

The simulation steps involved are as follows:

Step 1: The value of $x$ is generated from the stated distribution (normal, positively skewed, negatively skewed).

Step 2: Fit the multinomial logit equation:

$g_1(x) = \beta_{10} + \beta_{11} x_1 + \beta_{12} x_2$

$g_2(x) = \beta_{20} + \beta_{21} x_1 + \beta_{22} x_2$
Step 3: Evaluate the multinomial logistic probabilities for each category ($\pi_0, \pi_1, \pi_2$).

\[
\begin{align*}
\pi_0(x) &= \frac{1}{1 + e^{\gamma_1(x)} + e^{\gamma_2(x)}} \\
\pi_1(x) &= \frac{e^{\gamma_1(x)}}{1 + e^{\gamma_1(x)} + e^{\gamma_2(x)}} \\
\pi_2(x) &= \frac{e^{\gamma_2(x)}}{1 + e^{\gamma_1(x)} + e^{\gamma_2(x)}}
\end{align*}
\]

Step 4: The value of $U$ is generated by using an independent $U(0,1)$ distribution.

Step 5: Variable ($Y$) outcome is assigned based on the rule:

i. $Y=1$ if $u > \pi_0 + \pi_1$
ii. $Y=2$ if $u < \pi_0 + \pi_1$ and $u > \pi_0$
iii. $Y=0$ otherwise

Step 6: Fit multinomial logistic regression model based on the generated $x$ and $y$ and estimate $\beta$.

Step 7: Repeat Steps (1) – (6) for 10000 replications to obtain $\hat{\beta}^1, \hat{\beta}^2, \ldots, \hat{\beta}^{10000}$.

Step 8: Calculate the mean of $\hat{\beta}$, $\bar{\beta} = \frac{\sum_{i=1}^{10000} \hat{\beta}}{10000}$

Step 9: Calculate the 95% Confidence Interval (CI) for $\bar{\beta}$.

This simulation study involved three types of continuous distribution. N(0,1) represents normal distribution, Beta(12,1) represents negatively skewed and $\chi^2(4)$ represents positively skewed distribution.

3. Simulation Results

Simulation study is carried out in order to investigate the effect of covariate distribution on parameter estimates for multinomial logistic regression. Table 2 shows the summary of the parameter estimated and the 95% confidence interval (CI) for model that contains two normal covariate distribution (Model A). At sample size of 300 and above, the estimated parameter can be considered as closed to the true parameter value. Overall, the estimated parameters not deviate too far from the true parameter even the sample size used is small.
he estimated parameters \( \beta \) for the model that contains positively skewed covariate. The next, negatively skewed covariate requires larger sample size to get closer to the true parameter value at sample size of 1000 and above. This shows that covariate distributions which are N(0,1) and Beta(12,1). The value of parameter estimates get closer to the true parameter value at sample size of 1000 and above. This shows that the model that contains negatively skewed covariate requires larger sample size to get closer to the true parameter value compared to the model without negatively skewed distribution covariate. The estimated parameters become unbiased when the sample size is very large.

### Table 2. Parameter Estimation for Model A

| Sample size | \( \beta_{10} \) | \( \beta_{11} \) | \( \beta_{12} \) | \( \beta_{20} \) | \( \beta_{21} \) | \( \beta_{22} \) |
|-------------|------------------|------------------|------------------|------------------|------------------|------------------|
| 50          | -2.068           | -0.428           | 1.301            | -2.314           | -0.255           | 2.162            |
| 100         | -2.239           | -0.379           | 1.163            | -2.036           | -0.230           | 1.837            |
| 150         | -2.210           | -0.377           | 1.137            | -1.987           | -0.221           | 1.786            |
| 300         | -2.151           | -0.358           | 1.101            | -1.946           | -0.218           | 1.738            |
| 500         | -2.130           | -0.355           | 1.096            | -1.924           | -0.215           | 1.719            |
| 1000        | -2.114           | -0.353           | 1.087            | -1.912           | -0.213           | 1.703            |
| 1500        | -2.109           | -0.351           | 1.084            | -1.909           | -0.210           | 1.699            |
| 3000        | -2.105           | -0.351           | 1.083            | -1.904           | -0.210           | 1.694            |
| 5000        | -2.103           | -0.352           | 1.082            | -1.902           | -0.210           | 1.692            |

Table 3 shows the summary of parameter estimates and the 95% (CI) for Model B – model with two covariate distributions which are N(0,1) and Beta(12,1). The value of parameter estimates get closer to true parameter value at sample size of 1000 and above. This shows that the model that contains negatively skewed covariate requires larger sample size to get closer to the true parameter value compared to the model without negatively skewed distribution covariate. The estimated parameters become unbiased when the sample size is very large.

### Table 3. Parameter Estimation for Model B

| Sample size | \( \beta_{10} \) | \( \beta_{11} \) | \( \beta_{12} \) | \( \beta_{20} \) | \( \beta_{21} \) | \( \beta_{22} \) |
|-------------|------------------|------------------|------------------|------------------|------------------|------------------|
| 50          | -3.454           | -0.407           | 2.436            | -2.483           | -0.238           | 2.287            |
| 100         | -2.725           | -0.375           | 1.671            | -2.088           | -0.223           | 1.869            |
| 150         | -2.492           | -0.364           | 1.456            | -2.084           | -0.217           | 1.879            |
| 300         | -2.278           | -0.361           | 1.249            | -1.935           | -0.214           | 1.770            |
| 500         | -2.236           | -0.353           | 1.214            | -1.930           | -0.212           | 1.720            |
| 1000        | -2.138           | -0.355           | 1.113            | -1.913           | -0.211           | 1.702            |
| 1500        | -2.134           | -0.350           | 1.113            | -1.924           | -0.208           | 1.714            |
| 3000        | -2.118           | -0.351           | 1.096            | -1.912           | -0.211           | 1.702            |
| 5000        | -2.115           | -0.351           | 1.095            | -1.901           | -0.210           | 1.690            |

Next, Table 4 shows the summary of the parameter estimates and 95% CI for Model C, which is the model that contains positively skewed covariate. The result shows that the parameter estimates can be considered as close to the true parameter value at sample size of 300 and above. This indicates that the model that contains positively skewed covariate are less affected compared to the model that contains negatively skewed covariate.
Table 4. Parameter Estimation for Model C

| Sample size | $\beta_{10} = -2.10$ | $\beta_{11} = -0.35$ | $\beta_{12} = 1.08$ | $\beta_{20} = -1.90$ | $\beta_{21} = -0.21$ | $\beta_{22} = 1.69$ |
|-------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|
| 50          | 3.305                | -0.998               | 2.239                | -3.776               | -0.421               | 3.050                |
| 100         | -2.365               | -0.416               | 1.299                | -2.293               | -0.256               | 2.013                |
| 150         | (2.396-2.334)        | (0.432-0.399)        | (1.274-1.323)        | (2.323-2.262)        | (2.272-2.241)        | (1.988-2.037)        |
| 300         | -2.274               | -0.376               | 1.201                | -2.131               | -0.236               | 1.875                |
| 500         | -2.180               | -0.362               | 1.132                | -2.003               | -0.220               | 1.771                |
| 1000        | -2.148               | -0.364               | 1.113                | -1.965               | -0.220               | 1.742                |
| 1500        | (2.156-2.139)        | (0.368-0.360)        | (1.109-1.118)        | (1.973-1.958)        | (1.784-2.071)        | 1.713                |
| 3000        | -2.118               | -0.358               | 1.094                | -1.928               | -0.215               | 1.706                |
| 5000        | (2.124-2.112)        | (0.361-0.355)        | (1.091-1.097)        | (1.933-1.922)        | (1.710-1.716)        | (1.695-1.699)        |

Table 5 shows the summary of the parameter estimates for each model in all three distributions for comparison purposes. Based on this table, it can be seen clearly that Model B (model with negatively skewed covariate) needs larger sample size to get closer to the true parameter value compared to other models.

Table 5. Parameter Estimates for Different Distribution

| Sample size | Model | $\beta_{10} = -2.10$ | $\beta_{11} = -0.35$ | $\beta_{12} = 1.08$ | $\beta_{20} = -1.90$ | $\beta_{21} = -0.21$ | $\beta_{22} = 1.69$ |
|-------------|-------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|
| 50          | A     | -2.068               | -0.428               | 1.301                | -2.314               | -0.255               | 2.162                |
|             | B     | -3.454               | -0.407               | 2.436                | -2.483               | -0.238               | 2.287                |
|             | C     | -3.305               | -0.598               | 2.239                | -3.776               | -0.421               | 3.050                |
| 100         | A     | -2.239               | -0.379               | 1.163                | -2.036               | -0.230               | 1.837                |
|             | B     | -2.725               | -0.375               | 1.671                | -2.088               | -0.223               | 1.869                |
|             | C     | -2.365               | -0.416               | 1.299                | -2.293               | -0.256               | 2.013                |
| 150         | A     | -2.210               | -0.377               | 1.137                | -1.987               | -0.221               | 1.786                |
|             | B     | -2.492               | -0.364               | 1.456                | -2.084               | -0.217               | 1.879                |
|             | C     | -2.274               | -0.376               | 1.201                | -2.131               | -0.236               | 1.875                |
| 300         | A     | -2.278               | -0.361               | 1.249                | -1.935               | -0.214               | 1.770                |
|             | B     | -2.180               | -0.362               | 1.132                | -2.003               | -0.220               | 1.771                |
|             | C     | -2.130               | -0.355               | 1.096                | -1.924               | -0.215               | 1.719                |
| 500         | A     | -2.236               | -0.353               | 1.214                | -1.930               | -0.212               | 1.720                |
|             | B     | -2.148               | -0.364               | 1.113                | -1.965               | -0.220               | 1.742                |
|             | C     | -2.114               | -0.353               | 1.087                | -1.912               | -0.213               | 1.703                |
| 1000        | A     | -2.138               | -0.353               | 1.113                | -1.913               | -0.211               | 1.702                |
|             | B     | -2.188               | -0.358               | 1.094                | -1.928               | -0.215               | 1.713                |
|             | C     | -2.109               | -0.351               | 1.084                | -1.909               | -0.210               | 1.699                |
| 1500        | A     | -2.134               | -0.350               | 1.113                | -1.924               | -0.208               | 1.714                |
|             | B     | -2.114               | -0.353               | 1.090                | -1.921               | -0.212               | 1.706                |
|             | C     | -2.105               | -0.351               | 1.083                | -1.904               | -0.210               | 1.694                |
| 3000        | A     | -2.118               | -0.351               | 1.096                | -1.912               | -0.211               | 1.702                |
|             | B     | -2.108               | -0.352               | 1.085                | -1.908               | -0.212               | 1.697                |
|             | C     | -2.105               | -0.351               | 1.083                | -1.904               | -0.210               | 1.694                |
| 5000        | A     | -2.103               | -0.352               | 1.082                | -1.902               | -0.210               | 1.692                |
|             | B     | -2.115               | -0.351               | 1.095                | -1.901               | -0.210               | 1.690                |
|             | C     | -2.105               | -0.351               | 1.083                | -1.904               | -0.210               | 1.694                |

Model A: Normal (0, 1), Model B: Beta (12, 1), Model C: $\chi^2 (4)$
To measure how close are the fitted parameter estimate values, $\hat{\beta}$ to the true parameter values, the Mean Square Error (MSE) is used. Table 6 shows the MSE for Model A. The values of MSE for model A (both normal distributed covariate) become smaller (less than 1) at sample size of 100 and above.

| Sample size | $\beta_{10} = -2.10$ | $\beta_{11} = -0.35$ | $\beta_{12} = 1.08$ | $\beta_{20} = -1.90$ | $\beta_{21} = -0.21$ | $\beta_{22} = 1.69$ |
|-------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|
| 50          | 0.375                | 0.425                | 0.69                 | 23.341               | 1.985                | 16.008               |
| 100         | 0.235                | 0.17                 | 0.26                 | 0.237               | 0.129                | 0.275               |
| 150         | 0.157                | 0.106                | 0.161                | 0.136               | 0.074                | 0.154               |
| 300         | 0.061                | 0.047                | 0.069                | 0.059               | 0.034                | 0.067               |
| 500         | 0.035                | 0.027                | 0.04                 | 0.033               | 0.02                 | 0.038               |
| 1000        | 0.016                | 0.013                | 0.019                | 0.016               | 0.01                 | 0.018               |
| 1500        | 0.011                | 0.009                | 0.013                | 0.01                 | 0.007                | 0.012               |
| 3000        | 0.005                | 0.004                | 0.006                | 0.005               | 0.003                | 0.006               |
| 5000        | 0.003                | 0.003                | 0.004                | 0.003               | 0.002                | 0.003               |

Table 7 summarizes the MSE result for Model B (contains one negatively skewed covariate). The result shows that the MSE value is very large when the sample size is small. As sample size increases, the MSE value decreases. The model needs larger sample size to get smaller MSE value compared to the model without negatively skewed covariate distribution.

| Sample size | $\beta_{10} = -2.10$ | $\beta_{11} = -0.35$ | $\beta_{12} = 1.08$ | $\beta_{20} = -1.90$ | $\beta_{21} = -0.21$ | $\beta_{22} = 1.69$ |
|-------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|
| 50          | 65.037               | 0.274                | 73.089               | 31.665               | 0.151                | 36.288               |
| 100         | 22.057               | 0.110                | 25.154               | 11.904               | 0.061                | 13.715               |
| 150         | 12.484               | 0.066                | 14.36                | 7.099                | 0.22                 | 20.904               |
| 300         | 5.366                | 0.031                | 6.217                | 3.243                | 0.018                | 3.769               |
| 500         | 3.131                | 0.018                | 3.632                | 1.832                | 0.011                | 2.128               |
| 1000        | 1.008                | 1.710                | 1.400                | 55.606               | 19.490               | 17.927               |
| 1500        | 0.984                | 0.006                | 1.146                | 0.612                | 0.003                | 0.710               |
| 3000        | 0.475                | 0.003                | 0.549                | 0.297                | 0.002                | 0.346               |
| 5000        | 0.280                | 0.002                | 0.326                | 0.177                | 0.001                | 0.206               |

The values of MSE for Model C are summarized in Table 8. The values of MSE become smaller (less than 1) at sample size of 150 and above.
Table 8. MSE for Model C

| Sample size | $\beta_{10} = -2.10$ | $\beta_{11} = -0.35$ | $\beta_{12} = 1.08$ | $\beta_{20} = -1.90$ | $\beta_{21} = -0.21$ | $\beta_{22} = 1.69$ |
|-------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|
| 50          | 63.086               | 7.654                | 45.130               | 65.702               | 7.475                | 45.675               |
| 100         | 2.592                | 0.706                | 1.665                | 2.538                | 0.614                | 1.695                |
| 150         | 0.843                | 0.195                | 0.302                | 0.743                | 0.139                | 0.300                |
| 300         | 0.350                | 0.082                | 0.11                 | 0.287                | 0.057                | 0.103                |
| 500         | 0.194                | 0.047                | 0.06                 | 0.161                | 0.033                | 0.057                |
| 1000        | 0.094                | 0.023                | 0.028                | 0.076                | 0.016                | 0.026                |
| 1500        | 0.061                | 0.015                | 0.018                | 0.049                | 0.010                | 0.016                |
| 3000        | 0.030                | 0.007                | 0.009                | 0.024                | 0.005                | 0.008                |
| 5000        | 0.018                | 0.004                | 0.005                | 0.015                | 0.003                | 0.005                |

Table 9 summarizes the MSE values for all three models. The results show that the MSE values for the model that contains negatively skewed covariate are the largest compared to the other models. As expected, for all models, as the sample size is increases, the MSE value also decreases.

Table 9. MSE for Model A, B and C

| Sample size | Model | $\beta_{10} = -2.10$ | $\beta_{11} = -0.35$ | $\beta_{12} = 1.08$ | $\beta_{20} = -1.90$ | $\beta_{21} = -0.21$ | $\beta_{22} = 1.69$ |
|-------------|-------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|
| 50          | A     | 0.375                | 0.425                | 0.69                 | 23.341               | 1.985                | 16.008               |
|             | B     | 65.037               | 0.274                | 73.089               | 31.665               | 0.151                | 36.288               |
|             | C     | 63.086               | 7.654                | 45.130               | 65.702               | 7.475                | 45.675               |
| 100         | A     | 0.235                | 0.17                 | 0.26                 | 0.237                | 0.129                | 0.275                |
|             | B     | 22.057               | 0.110                | 25.154               | 11.904               | 0.061                | 13.715               |
|             | C     | 2.592                | 0.706                | 1.665                | 2.538                | 0.641                | 1.695                |
| 150         | A     | 0.157                | 0.106                | 0.161                | 0.136                | 0.074                | 0.154                |
|             | B     | 12.484               | 0.066                | 14.36                | 7.099                | 0.220                | 20.904               |
|             | C     | 0.843                | 0.195                | 0.302                | 0.743                | 0.139                | 0.300                |
| 300         | A     | 0.061                | 0.047                | 0.069                | 0.059                | 0.034                | 0.067                |
|             | B     | 5.366                | 0.031                | 6.217                | 3.243                | 0.018                | 3.769                |
|             | C     | 0.350                | 0.082                | 0.11                 | 0.287                | 0.057                | 0.103                |
| 500         | A     | 0.035                | 0.027                | 0.04                 | 0.033                | 0.020                | 0.038                |
|             | B     | 3.131                | 0.018                | 3.632                | 1.832                | 0.011                | 2.128                |
|             | C     | 0.194                | 0.047                | 0.060                | 0.161                | 0.033                | 0.057                |
| 1000        | A     | 0.016                | 0.013                | 0.019                | 0.016                | 0.010                | 0.018                |
|             | B     | 1.008                | 1.710                | 1.400                | 55.606               | 19.49                | 17.927               |
|             | C     | 0.094                | 0.023                | 0.028                | 0.076                | 0.016                | 0.026                |
| 1500        | A     | 0.011                | 0.009                | 0.013                | 0.01                 | 0.007                | 0.012                |
|             | B     | 0.984                | 0.006                | 1.146                | 0.612                | 0.003                | 0.71                 |
|             | C     | 0.061                | 0.015                | 0.018                | 0.049                | 0.010                | 0.016                |
| 3000        | A     | 0.005                | 0.004                | 0.006                | 0.005                | 0.003                | 0.006                |
|             | B     | 0.475                | 0.003                | 0.549                | 0.297                | 0.002                | 0.346                |
|             | C     | 0.030                | 0.007                | 0.009                | 0.024                | 0.005                | 0.008                |
| 5000        | A     | 0.003                | 0.003                | 0.004                | 0.003                | 0.002                | 0.003                |
|             | B     | 0.280                | 0.002                | 0.326                | 0.177                | 0.001                | 0.206                |
|             | C     | 0.018                | 0.004                | 0.005                | 0.015                | 0.003                | 0.005                |

4. Conclusion
MLR is widely applied in many research fields such as medical, business, finance, computer science and environmental research. The aims of this study is to evaluate the effect of different distribution of
covariates on parameter estimation for multinomial logistic regression model and to determine the minimum sample size required to minimize the effect of covariate distribution on parameter estimation for multinomial logistic regression model. The simulation result shows that the parameter estimation are affected by the distribution of covariate. A large sample size is required to produce unbiased estimates for the model that contains skewed covariate distribution compared to the model that contains no skewed covariate distribution. The parameter estimation is less affected for model that contains only normal covariate distribution and the model that contains positively skewed covariate distribution compared to the model that contains negatively skewed covariate distribution. For model that contains only normal covariate distribution, and model that contains positively skewed covariate distribution, the parameter estimates can be considered as close to the true parameter value when the sample is 300 and above, while for the model that contains negatively skewed covariate distribution, it needs sample size of 1000 and above. Overall, the results from this simulation study is consistent with the results in Hamid et al. (2016), which stated that the MLR model with only negatively skewed covariate distribution is more affected compared to the model with only normal and model with only positively skewed covariate distribution. Future research can be carried out to investigate the effect of count and categorical data on MLR. More information also can be gathered if the covariates are mixed between continuous, categorical and count data.

References
[1] Khan, A., & Rayner, G. D. (2004). Robustness to non-normality of common tests for the many-sample location problem. Journal of Applied Mathematics and Decision Sciences, 7(4), 187–206.
[2] Jahan, S., & Khan, A. (2012). Power of t-test for simple linear regression model with non-normal error distribution: a quantile function distribution approach. Journal of Scientific Research, 4(3), 609–622.
[3] Hamid, A. H., Yap, B. W., & Xie, X.-J. (2016). Effects of different type of covariates and sample size on parameter estimation for multinomial logistic regression model, Jurnal Teknologi UTM(3), 155–161.
[4] Hosmer Jr, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). Applied logistic regression (Vol. 398). (3rd Edition). (Vol. 398). John Wiley & Sons.
[5] Fagerland, M. W., Hosmer, D. W., & Bofin, A. M. (2008). Multinomial goodness-of-fit tests for logistic regression models. Statistics in Medicine, 27(21), 4238–4253.
[6] Hamid A. H., Yap, B. W., Xie, X.-J., & Hezlin A.R. (2015). Assessing the Effects of Different Types of Covariates for Binary Logistic Regression. AIP Conference Proceedings 1643, 425–430.
[7] Kutner, Michael H., Christopher J. Nachtsheim, John Neter, W. L. (2005). Applied Linear Statistical Models (Fifth Edition). McGraw Hill.
[8] McCullagh, P., & Nelder, J. A. (2007). Generalized linear models (Vol.37) CRC press.
[9] Xie, X. J., Pendergast, J., & Clarke, W. (2008). Increasing the power: A practical approach to goodness-of-fit test for logistic regression models with continuous predictors. Computational Statistics and Data Analysis, 52(5), 2703–2713.