

Finding Factors Caused Diabetes Mellitus Disease Using Link Functions

¹Sudarno, ²Tatik Widiharih, ³Moch. Abdul Mukid

{dsghani@gmail.com}

^{1,2,3}Department of Statistics, Faculty Sciences and Mathematics,
Diponegoro University, Semarang, Indonesia

Abstract. Diabetes mellitus disease was a disease caused abnormal of pancreas organ. Abnormal of pancreas organ causes supply insulin in the blood be a few. If a person has diabetes mellitus for long time, then that body would have complication with other disease or body organ. This disease could be known by level of blood sugar after fast. If a person has symptom level of blood sugar after fast was greater than 126 mg/dL, he could be said that he has diabetes mellitus disease. Many people have diabetes mellitus in this time. Generally, a person could get this disease, because genetic from parents, human life about eat and drink, and bad hobby. Link function was function to relate between response variable versus predictor variables in multiple linear regression. Predictor variables of link function were linear with respect to both variable and parameter. The link functions of this paper were logit link function, normit link function and complementary log-log (cloglog) link function. We take response variable is diabetes mellitus, while the factors cause this disease (covariate variables) are age, sex, systolic blood pressure, diastolic blood pressure, logarithm of the ratio of urinary albumin and creatinine, high-density lipoprotein in cholesterol, logarithm of insulin, and smoke. This paper wants to determine the factors cause diabetes mellitus by significant. In this research, the significant factors were sex, logarithm of the ratio of urinary albumin and creatinine, and logarithm of insulin. If the factors cause it be found, so its disease could be prevented toward illness of persons. The significant factors could be made the best model by three link functions, respectively.. The best model could be used to predict probability of a person has diabetes mellitus disease. Then this probability could be used to predict chance a person with diabetes mellitus disease. The linear hypothesis test be used to measure significant of differences about two link functions. This result could be said that there are no significant of differences about two link functions. So three link functions to predict chance a person has diabetes mellitus obtained almost same values. The goodness of fit model by AIC value. AIC value of logit link function, normit link function, and cloglog link function, respectively, 52.778, 52.639, 52.257. These results are same with the result of linear hypothesis test that AIC values are not different. Characteristics relation between link functions that estimated probability, and AIC value of link functions, respectively, are not significant or in the same cluster. Therefore, logit model, normit model and cloglog model could be used to predict probability with same result.

Keywords: Diabetes mellitus, link function, logit, normit, cloglog, AIC, linear hypothesis test.

1. Introduction

Diabetes mellitus disease was disease which abnormal metabolism for long time, because pancreas could not be able to produce insulin hormone be enough, or because that body could not be able to use insulin hormone had been produced by effective. Insulin hormone was hormone its function for regulate stable level of blood sugar in body. If pancreas function was bad and insulin hormone was a few, then the body would have high level of blood sugar (hiperglicemia). This condition could cause abnormal to many organs of patient body. The body was set of many organs, they have function themselves, but they have fixed task. If level of blood sugar was not controlled by well, it could cause many various complication body organs. The diseases often appear at patient with high level of blood sugar were coroner heart, stroke, high blood pressure, and kidney. Generally, a person could get this disease, because genetic from parents, human life about eat and drink, and bad hobby (WHO, 2010).

There are some symptoms or way to detection about a person has diabetes mellitus. Diabetes disease symptoms often appear as often thirsty, urinate (poliuria), weak, and easy infection. Way to detect a person has this disease could use level of sugar in urine or blood after fast. If level of blood sugar after fast less than or equal to 126 mg/dL, then a person could not have diabetes mellitus disease. But, if a person has level of blood sugar after fast greater than 126 mg/dL, then this person has symptom of diabetes mellitus disease. There are two types of diabetes mellitus, that are type 1 and type 2. Type 2 was diabetes disease which a body still produce insulin hormone, but a few amount. Factors cause this disease are genetic or overweight. This disease could be done by regulate level of insulin in blood. The way by inject insulin hormone at scheduled time (Rotella et al., 2015).

Regression is formula constructed between independent variable and dependent variable. Formally, regression could be defined as relation between response variable versus predictor variables. Response variabel could have categoric and interval or rasio type, then predictor variables could be measured by nominal or ordinal scale, and the others have interval or ratio scales. Regression has response variabel was categorical, it could be divided into two groups. The groups are response variable with two categories and response variable with more than two categories. Response variable has two categories named by response variable of binary, while response variable has more than two categories named by response variable of multinomial^[1, 5]. In Lee et al. (2013) said that binary response variable could be made as binary regression. Binary regression was regression it has response variable yield binary outcomes. Binary outcomes be able to be produced from predictor variable with data measured by arbitrary scale.

Link function was function of multiple linear regression, that was linear toward both predictor variable and parameter. Link functions in this research were logit link function, normit link function and cloglog link function. Logit link function has domain $[0, 1]$ and range $(-\infty, \infty)$, while normit function is quantile function related with normal standard distribution. The normal standart distribution was normal distribution with mean 0 and standard deviation 1. Mathematically, normit function was reciprocal of cumulative distribution function of normal standard distribution^[2, 7, 8].

Some journals related with this paper that are in Rozzaghi (2013) discuss about logistic regression model be compared with probit regression model applied at data mining. Logistic regression model initially come from logit transformation, while probit regression model initially come from probit transformation. The result that logit transformation can be collected

to class of canonic link function, but probit transformation can not be collected to class of canonic link function. Then at Cakmayapan, et al., (2013) explained that comparison logit regression model and probit regression model by Monte Carlo simulation with different sample size, different cutoff point, and different correlation between dependent variable and independent variables. This material discuss models, residuals, deviations, and Pseudo-R square. Comparison link function with binary response variable about symmetric assumption that probit link function be better than logit link function. Because value of AIC of probit link function be smaller than value of AIC of logit link function. While for loglog link function be better than cloglog link function. Because value of AIC of loglog link function be smaller than value of AIC of cloglog link function (Damisa, et al. 2017). At Shariff, et al. (2009) that logit regression model more robust than probit regression model for data were not normal distribution.

This paper explain about find factors caused diabetes mellitus disease by logit link function, normit link function, and cloglog link function. It find the best model be resulted link function, respectively. Then they be applied to compute probability values of a person has disease diabetes mellitus. Test of homogeneity prediction result of link functions use linear hypothesis test. These factors cause diabetes mellitus could be used to prevent a person in order to be avoided from diabetes mellitus disease.

2. Method and Material

This paper was application research. Therefore, the method was presented by simple and short in order to be applicable. This section contains method and material of research. They were hoped to be able answer the result and discussion section. We used R software and statistics program by survival to process these data.

2.1 Logistic Regression Model of Binary Outcomes

Let there were n patients observed in research. Then patient status was collected by two groups, that were ill and healthy groups. If it was ill equal yes, while if it was healthy equal no. So, there were two categories about patients had a disease. Let $y_i = 1$ was patient has a disease (yes), and $y_i = 0$ was patient has not a disease (no), for $i = 1, 2, \dots, n$. The other ones, they were measured p independent variables $x_{i1}, x_{i2}, \dots, x_{ip}$, that were covariate variables of every patient. The covariate variables could be qualitative and quantitative variables. The example of qualitative variables was sex, race, while for quantitative variables were blood pressure, amount of leucosite cells. This paper wanted to look for relationship between independent variables $x_{i1}, x_{i2}, \dots, x_{ip}$ with dependent variable binary $y_i, i = 1, 2, \dots, n$ ^[1, 9].

Let P_i was probability of yes, means that $P_i = P(y_i = 1 | x_{i1}, \dots, x_{ip})$ for i -th patient. Then by Cox (1970) proposed that logistic regression model about relation probability of yes with independent variable could be written by

$$P_i = P(y_i = 1 | \mathbf{x}_i) = \frac{\exp\left(\sum_{j=0}^p b_j x_{ij}\right)}{1 + \exp\left(\sum_{j=0}^p b_j x_{ij}\right)}$$

(1)

and

$$1 - P_i = P(y_i = 0 | \mathbf{x}_i) = \frac{1}{1 + \exp\left(\sum_{j=0}^p b_j x_{ij}\right)}$$

(2)

where $\mathbf{x}_i = (x_{i0}, \dots, x_{ip})$, $x_{i0} \equiv 1$, and $b_j, j = 0, 1, \dots, p$ were unknown coefficients. If this is formulated as

$$\frac{P_i}{1 - P_i} = \exp\left(\sum_{j=1}^p b_j x_{ij}\right)$$

Then that equation by logarithm be obtained relation that was

$$\lambda_i = \log\left(\frac{P_i}{1 - P_i}\right) = \sum_{j=1}^p b_j x_{ij}$$

This equation was a multiple linear regression. Generally, this equation be said linear logistic regression model. This model often be used for data with binary response and covariate variables with normal distribution^[1, 9]. For to estimate coefficient $b_j, j = 0, 1, \dots, p$ by Cox (1970) recommended by maximum likelihood estimation method. This method was as follow. If $y_i, i = 1, 2, \dots, n$ be observations that have binary values at n patients, then these variables have binomial distribution which formed by (1) whenever $y_i = 1$ and (2) whenever $y_i = 0$. Thus, the likelihood function was

$$\begin{aligned} L(b_0, b_1, \dots, b_p) &= \prod_{i=1}^n P_i^{y_i} (1 - P_i)^{1-y_i} \\ &= \frac{\exp\left(\sum_{j=0}^p b_j t_j\right)}{\prod_{i=1}^n \left[1 + \exp\left(\sum_{j=0}^p b_j x_{ij}\right)\right]} \end{aligned}$$

where $t_j = \sum_{i=1}^n x_{ij} y_j$. The log-likelihood function was

$$l(b_0, b_1, \dots, b_p) = \log L = \sum_{j=0}^p b_j t_j - \sum_{i=1}^n \log \left[1 + \exp \left(\sum_{j=0}^p b_j x_{ij} \right) \right]$$

Then be looked for parsial derivative of $l(b_0, b_1, \dots, b_p)$ with respect to repectively parameters and they were equal zero. This process has $(p + 1)$ equations simultaneously. If those equations be written by short obtained

$$t_k - \sum_{i=1}^n \frac{x_{ik} \exp \left(\sum_{j=0}^p b_j x_{ij} \right)}{1 + \exp \left(\sum_{j=0}^p b_j x_{ij} \right)} = 0 \quad k = 0, 1, \dots, p.$$

These equations could be solved by Newton-Raphson procedure. To proof that their parameter values maximize log-likelihood function, then it needs to look for second parsial derivative of $l(b_0, b_1, \dots, b_p)$ with respect to parameters, respectively, where the result was negatif, that was

$$I_{j_1 j_2}^* = \frac{\partial^2 l}{\partial b_{j_1} \partial b_{j_2}} = - \sum_{i=1}^n \frac{x_{ij_1} x_{ij_2} \exp \left(\sum_{j=0}^p b_j x_{ij} \right)}{1 + \exp \left(\sum_{j=0}^p b_j x_{ij} \right)}, \quad j_1 = 0, \dots, p; j_2 = 0, \dots, p$$

2.2 Various Link Functions

Link function was function which defined the relationship between the covariate and probability of succes. Generally, this function could be explained that if there was pobability of succes of event i -th, that was P_i , with covariate variables amount p , x_1, x_2, \dots, x_p , then link function was response variable of function P_i , and the predictor variables built linear combination of p -covariates. For a probability of succes from individual i -th, P_i , then link function could be presented by $g(P_i)$ [7, 8, 9, 13].

Formally, there were three types of link functions, that were

- a. Logit link function

Logit link function was defined by logit function, that was

$$g(P_i) = \log \left(\frac{P_i}{1 - P_i} \right)$$

Then link function was

$$\log \left(\frac{P_i}{1 - P_i} \right) = \sum_{j=0}^p b_j x_{ij} \quad (3)$$

This function be used as response variable which has binary outcomes. If equation (3) was extended by property of logarithm with respect to variable P_i , then we be obtained function

$$P_i = \frac{\exp\left(\sum_{j=0}^p b_j x_{ij}\right)}{1 + \exp\left(\sum_{j=0}^b b_j x_{ij}\right)} \quad (4)$$

Equation (4) could be used to find probability value of individual i -th if the covariate values were known.

b. Normit link function

Normit link function was defined to use inverse of cumulative standard normal distribution function, $\Phi^{-1}(\cdot)$, that was

$$g(P_i) = \Phi^{-1}(P_i)$$

Therefore, the normit link function was

$$\Phi^{-1}(P_i) = \sum_{j=0}^p b_j x_{ij} \quad (5)$$

This link function should be used for respose variable has binary outcomes. Response variable binary outcomes be obtained from symmetric distribution data cut by cutoff point. The example of symmetric distribution, such as normal distribution. If equation (5) be made composite function by its inverse, then it obtained function as

$$P_i = \Phi\left(\sum_{j=0}^p b_j x_{ij}\right) \quad (6)$$

The equation (6) could be used to determine probability of individual i -th by cumulative standard normal distribution function towards its covariate value.

c. Complementary log-log link function

Complementary log-log link function was defined to use based on log-log function, that was

$$g(P_i) = \log(-\log(1 - P_i))$$

Therefore, the link function was

$$\log(-\log((1 - P_i))) = \sum_{j=0}^p b_j x_{ij} \quad (7)$$

This link function could be used to respose variable has binary outcomes. Binary outcomes be obtained from asymmetric distribution data cut by cutoff point. Example of asymmetric distribution was chi-square or F distribution. If equation (7) transformed by natural logaritm function by exponential function, then be obtained function

$$P_i = 1 - \exp\left(-\exp\left(\sum_{j=0}^p b_j x_{ij}\right)\right)$$

(8)

Equation (8) could be used to find probability of individual i -th when value of covariate variables be known.

2.3 Fitting model for prediction

Akaike information criterion (AIC) (Akaike, 1969), was defined as

$$AIC = l(\hat{\mathbf{b}}) - 2p$$

where

$l(\hat{\mathbf{b}})$: loglikelihood of β prediction

p : the number of parameters in distribution

Criterion of AIC be used to measure exactly of choosing the best regression model. The way define AIC values by formula that was smaller AIC value be better to choose the best regression model^[3].

2.4 Data and Research Variables

These data were secondary data about patient of diabetes mellitus disease. Souce of data from medical record section of Tugu hospital Semarang, Central Java, on February 2019. The research variables consist of age, sex, systolic blood pressure (sbp), diastolic blood pressure (dbp), logarithm of the ratio of urinary albumin and creatinine (lacr), high-density lipoprotein in cholesterol (hdl), logaritm of insulin (linsul), smoke, diabetes mellitus (dm). Respose variabel was dm, while covariate variables were age, sex, sbp, dbp, lacr, hdl, linsul, and smoke. The age variabel (year); sex = 1 if male and = 0 if female; sbp (mmHg); dbp (mmHg); smoke = 1 if current smoker and 0 otherwise; dm =1 if fasting glucose \geq 126 mg/dL and = 0 otherwise.

3. Result and Discussion

Application of this research was health field about a person has diabetes mellitus disease. The objective of discussion was how to know some factors which effect to get the diabetes mellitus disease. Its discussion use statistical concept. Method of analysis used logit link function, normit link function, and cloglog link function.

3.1 Logit Link Function Model

This research taken covariate variables and one response variable. The covariate variables consist of age, sex, sbp, dbp, lacr, hdl, linsul, and smoke. The response variabel was dm. Method of research process by regression between response variable versus some covariate variables. Every process found covariate variables were significant. The step of process by eliminating covariate variable based to value of $\Pr(>|z|)$ that highest. Final process obtained some the best significant covariate variables, that were sex, lacr, and linsul. Therefore, this research could be claimed that the factors caused diabetes mellitus disease were sex, lacr and linsul. In this research for sex variable that female total had diabetes mellitus greater than male total. So we can said that female had more risk than male about risk diabetes mellitus disease. The other ones, amount of albumin and creatinine in urine should be

controlled regularly. Level of insulin in blood must be enough, because insulin had task to cauch sugar in order to be meat.

Table 1. Covariate coefficient used logit function

Covariate	Estimate	Std. Error	z-value	Pr(> z)
Intercept	-11.2628	3.6887	-3.053	0.00226
sex	2.8760	1.3490	2.132	0.03301
lacr	1.5467	0.5275	2.932	0.00337
linsul	1.6427	0.8606	1.909	0.05629

While value of AIC = 52.778. Therefore, it could be made the best model from this process. The best logit model was

$$\log\left(\frac{P}{1-P}\right) = 2.8760 \text{ sex} + 1.5467 \text{ lacr} + 1.6427 \text{ linsul} - 11.2628$$

The best logit model would be applicated to some numbers of covariate variables, that were sex, lacr and linsul with result in the table 2 below.

Table 2. Estimation of probability values used logit function

I	sex	lacr	linsul	P
1	1	1.56	2.53	.1397
2	0	2.52	3.17	.1036
3	1	1.25	2.77	.1298
4	0	3.00	3.30	.2312
5	1	2.10	3.04	.4638

3.2 Normit Link Function Model

Based on process at link function model of normit have been obtained that covariate variables which significant are sex, lacr and linsul. Therefore, if the covarate variables be counted to normit link function obtained result as follow.

Table 3. Covariate coefficient used normit function

Covariate	Estimate	Std. Error	z-value	Pr(> z)
Intercept	-5.9297	1.8367	-3.228	0.00124
sex	1.4688	0.6962	2.110	0.03488
lacr	0.8628	0.2708	3.186	0.00144
linsul	0.8051	0.4576	1.759	0.07851

Where value of AIC = 52.639. Therefore, it could be made the best model from the normit link function.

The best normit model was

$$\Phi^{-1}(P) = 1.46881 \text{ sex} + 0.8628 \text{ lacr} + 0.8051 \text{ linsul} - 5.9297$$

This model be applied for couting some fixed numbers of covariate variables, that were sex, lacr and linsul. Computation result of this model was written at table 4, next section.

Table 4. Estimation of probability values by normit function

I	sex	lacr	linsul	P
1	1	1.56	2.53	.1405
2	0	2.52	3.17	.1144
3	1	1.25	2.77	.1246
4	0	3.00	3.30	.2468
5	1	2.10	3.04	.4201

3.3 Cloglog Link Function Model

The result of process analysis to link function models ago, it had been obtained that significant covariate variables were sex, lacr and linsul. If that covariate variables be applied to link function cloglog obtained result this following.

Table 5. Covariate coefficient using cloglog function

Covariate	Estimate	Std. Error	z-value	Pr(> z)
Intercept	-10.5644	3.1717	-3.331	0.000866
sex	2.6131	1.1583	2.256	0.024067
lacr	1.2943	0.4478	2.891	0.003846
linsul	1.6023	0.7219	2.220	0.026446

Where value of AIC = 52.257. The best model could be constructed from this table.

The best model of complementary log-log function was

$$\log(-\log(1 - P)) = 2.6131 \text{ sex} + 1.2943 \text{ lacr} + 1.6023 \text{ linsul} - 10.5644$$

Application of the best model of complementary log-log function was tried to count probability values of covariate variables. Some fixed numbers be taken from covariate variables: sex, lacr, and linsul, respectively, then the result was presented at table 6 as follow.

Table 6. Estimation of probability values using cloglog function

I	sex	lacr	linsul	P
1	1	1.56	2.53	.1417
2	0	2.52	3.17	.1026
3	1	1.25	2.77	.1396
4	0	3.00	3.30	.2197
5	1	2.10	3.04	.5015

Next discussion was finding properties of characteristics between three link functions with respect to covariate variables, comparative of probability values, and AIC values.

Relation between link functions with respect to covariate

If they were made relation between link functions to significant covariate variables, respectively, with respect to estimate and std. error values had expression that if the statistics values were ordered by small to big. They yielded the results be regularly. The ordering of link functions by normit link function, cloglog link function, and then probit link function, respectively. For explicitly these expressions were tabled as follow.

Table 7. Sex covariate variable

Link Function	Estimate	Std. Error
normit	1.4688	0.6962

cloglog	2.6131	1.1583
logit	2.8760	1.3490

Table 8. Lacr covariate variable

Link Function	Estimate	Std. Error
normit	0.8628	0.2708
cloglog	1.2943	0.4478
logit	1.5467	0.5275

Table 9. Linsul covariate variable

Link Function	Estimate	Std. Error
normit	0.8051	0.4576
cloglog	1.6023	0.7219
logit	1.6427	0.8606

But if we compare about intercept values of link functions respectively, these results were different sign. Those signs show values from big to small or decrease, and the order of link functions was same, respectively. The order of link functions were normit, cloglog, and then logit link functions.

Comparative of probability values

Linear hypothesis was used to test comparison of probability values by three link functions. The hypothesis was written as follow:

Hypothesis:

H_0 : all of link function methods result estimation probability value were same.

H_1 : minimum one of link function method result estimation probability value was different.

The output result was tabled below.

Table 10. Comparison between two methods

Method	Estimate	Std. Error	t-value	$\Pr(> t)$
B – A	-0.00434	0.09307	-0.047	0.999
C – A	0.00740	0.09307	0.080	0.997
C – B	0.01174	0.09307	0.126	0.991

If we were taken significant level $\alpha = 5\%$, then by $\Pr(>|t|)$ values. They could be concluded that H_0 was accepted. It means that all of link function methods yield estimation of probability value were same.

If above hypothesis was presented by graph, the figure as follow.

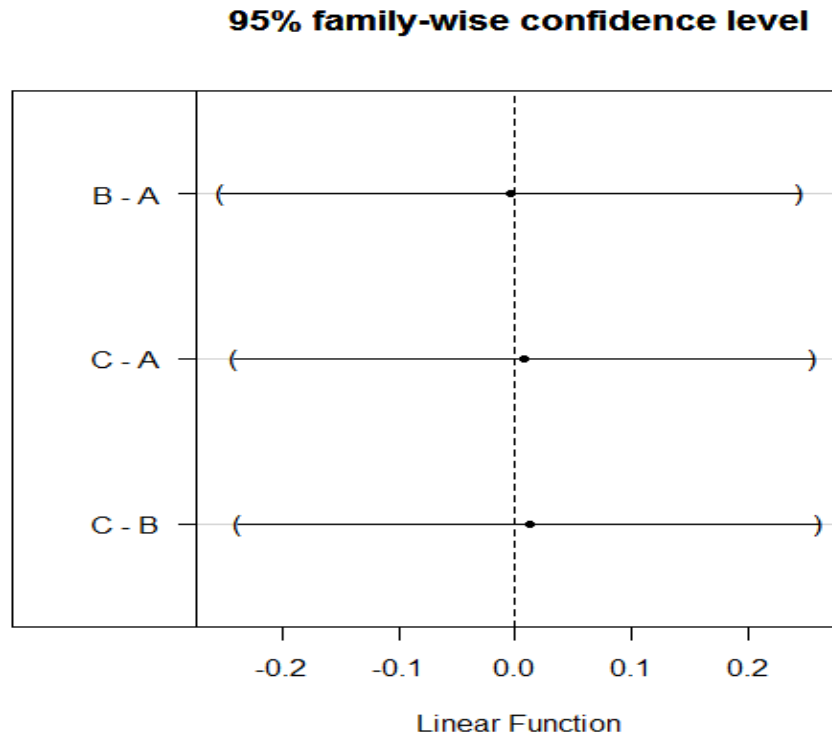


Figure 1. Differences between two link functions to predict probability.

AIC value

AIC value of logit link function, normit link function, and cloglog link function were, respectively, 52.778, 52.639, 52.257. These results were same with the result of linear hypothesis analysis that AIC values were not different while their AIC values were not equal. Therefore, logit model, normit model and cloglog model could be used to predict probability.

4. Conclusion

There were some important items to inform result of this paper.

1. A person could have diabetes mellitus disease be caused by the covariate variables that were sex, albumin level, degree of creatine in urine, and low level of insulin in blood. Female has more risk diabetes mellitus disease than male.
2. The best model of logit function was

$$\log\left(\frac{P}{1-P}\right) = 2.8760 \text{ sex} + 1.5467 \text{ lacr} + 1.6427 \text{ linsul} - 11.2628$$

3. The best model of normit function was

$$\Phi^{-1}(P) = 1.46881 \text{ sex} + 0.8628 \text{ lacr} + 0.8051 \text{ linsul} - 5.9297$$
4. The best model of cloglog function was

$$\log(-\log(1 - P)) = 2.6131 \text{ sex} + 1.2943 \text{ lacr} + 1.6023 \text{ linsul} - 10.5644$$

5. Characteristics relation between three link functions for estimate probability of diabetes mellitus disease, and AIC value of link functions, respectively, were not significant different or in the same cluster.

References

- [1] Agresti, A.: *An Introduction to Categorical Data Analysis*, Second Edition, John Wiley & Sons, Inc., New Jersey. (2007)
- [2] Agresti, A.: *Foundations of Linear and Generalized Linear Models*, John Wiley & Sons., New York. (2015)
- [3] Akaike, H.: Fitting Autoregressive Models for Prediction, *Annals of the Institute of Statistical Mathematics*. Vol. 21, pp. 243-247. (1969)
- [4] Cakmayapan, S., and Goktas, A.: A Comparison of Binary Logit and Probit Models with a Simulation Study, *Journal of Social and Economic Statistics*. Vol. 2 (1), Summer. (2013)
- [5] Cox, D.R.: *Analysis of Binary Data*, Methuen, London. (1970)
- [6] Damisa, S.,A., Bello, S.Y., and Musa, F.N.: On the Comparison of Some Link Functions of Binary Response Analysis Under Symetric and Asymetric Assumptions, *Biomedical Statistics and Informatics*. Vol. 2 (4), pp. 145-149. (2017)
- [7] Dobson, A.J.: *An Introduction to Generalized Linear Models*, Capman & Hall, London. (2001)
- [8] Hardin, J.W. and Hilbe, J.M.: *Generalized Linear Model and Extentions*, United State of America: Stata Corp LP. (2007)
- [9] Lee, E.T. and Wang, J.W.: *Statistical Methods for Survival Data Analysis*, Fourth Edition, John Wiley & Sons, Inc., New Jersey. (2013)
- [10] Rotella, C.M., Pola, L. and Mannuci, E.: Role of Insulin in the Thype 2 Diabetes Therapy, *International of Journal of Endocrinology Metabolism*. Vol. 11 (3), pp. 137-144. (2015)
- [11] Rozzaghi, M.: The Probit Link Function in Geneeralized Linear Models for Data Mining Applications, *Journal of Modern Applied Statistical Methods*, Bloomsberg University, Bloomsberg. Vol. 12 (1), pp. 164-169. (2013)
- [12] Shariff, A.A., Zaharim, A. and Sopian, K.: The Comparison of Logit and Probit Regression Analysis in Estimating the Strength of Gear Teeth, *European Journal of Scientific Research*. Vol. 27(4), pp. 548-553. (2009)
- [13] Tsay, R.S.: *Analysis of Financial Time Series*, Second Edition, ISBN-10: 0471690740, John Wiley & Sons, Inc., New Jersey. (2005)
- [14] World Health Organization Department of Noncommunicable Disease Surveilance: *Definition, Diagnosis and Classification of Diabetis Mellitus and Its Complication*, Switzerland. (2010)