

## LOGISTIC REGRESSION ANALYSIS IN PERSONAL LOAN BANKRUPTCY

Siti Mursyida Abdul Karim & Dr. Haliza Abdul Rahman

### Abstract

Personal loan bankruptcy is defined as a person who had been declared as a bankrupt due to failure to repay their personal loan. Personal loan bankruptcy is a serious problem that will affect an individual financial stability. This study focused on personal loan bankruptcy in Kedah only. For this study, the concept of logistic regression model was applied in order to determine the most predictive factor of personal loan bankruptcy problem. There are four main factors considered in this analysis, that is age, gender, race and job profession. Logistic regression emphasizes the nature of relationship between the dependent variable and another independent variables. The outcomes are predicted by using odd ratio. The odds ratio interpretation of the estimated regression coefficients makes the logistic regression model especially attractive for modelling and interpreting the studies. In this study, the data consist of 576 person declared as bankrupt due to personal loan bankruptcy and non-personal loan bankruptcy. The response variable is binary, denoting whether a person is personal loan bankruptcy or non-personal loan bankruptcy. The data are analysed by using SPSS 22. Based on the analysis, gender, race and job profession are the significant factors that lead to personal loan bankruptcy.

Keywords: Logistic regression, personal loan, bankruptcy

### Introduction

Regression methods have become an integral component of any data analysis concerned with describing the relationship between a response variable and one or more explanatory variables. Quite often the outcome variable and one is discrete, taking on two or more possible values. The logistic regression model is the most frequently used regression model for the analysis of these data.

Logistic regression is a type of generalized linear model that uses statistical analysis to predict an event based on known factors. It is also called as a logistic model and logit model. This broad class of models includes ordinary regression as well as multivariate statistics and loglinear regression. Logistic regression allows one to predict a discrete outcome, such as group membership, from a set of variables that may be continuous, discrete, dichotomous, or a mix of any these. Generally, the dependent or response variable is dichotomous, such as presence or absence and success or failure.

There are two main uses of logistic regression, the first is the prediction of group membership. Since logistic regression calculates the probability of success over the probability of failure, the results of the analysis are in the form of an odds ratio. In addition, logistic regression also provides knowledge of the relationships and strengths among the variables. Logistic regression is widely used in many field such as medical and financial field. For example, the uses of logistic regression is to evaluate whether bilirubin levels in cord blood could predict neonatal Hyperbilirubinemia that would require treatment, in full-term healthy new-born infants(Gatea,2009) .However, in financial field, it is used to predict of bankruptcy of small firms (Blanco ,2012).

## Literature Review

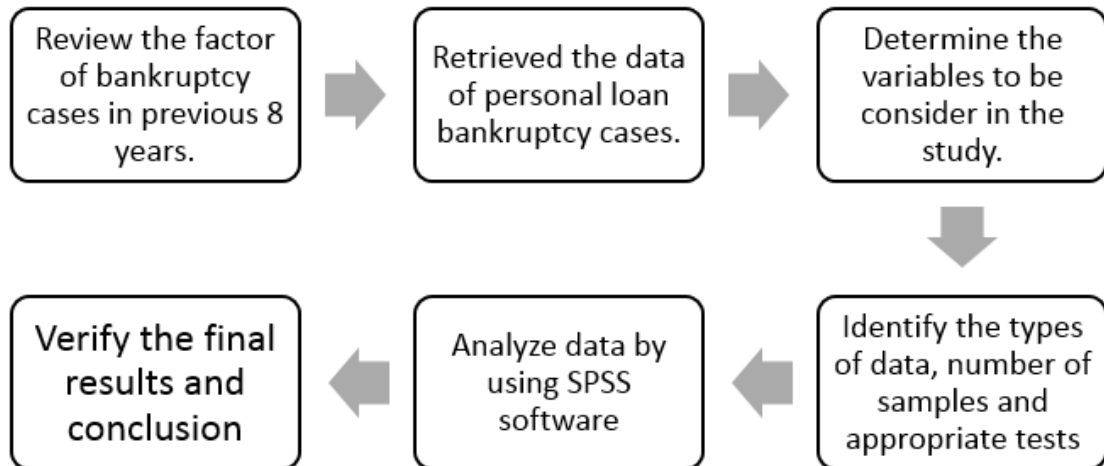
There is now an increment of bankruptcy in Malaysia with over 47% of young adults' age between 18 – 35-years-old are having serious debts due to living beyond their means. According to the recent statistics from the Malaysian Department of Insolvency (MDI), an average of 1,812 people are declared bankrupt per month last year, with an 11% rise in the average number of monthly bankruptcies from 2012 to 2013. One of the Unit Trust Consultant Manager, Erin said, the lack of discipline on financial management is why many young Malaysians today carrying major debts and also with it a high possibility to go bankrupt. Apart from majority of credit card debts and higher purchase loans like car and housing loans, most young adults did not settle their study loan as well upon graduating from university and this set a bad record or being blacklisted in Malaysia (Muzammil, 2014).

For the study conducted in order to determine whether non-performing loan, unemployment and economic condition affect the bankruptcy cases in Malaysia, based on the findings, both non-performing loan and unemployment have statistically significant relationship with bankruptcy. While, the unemployment is not significant towards bankruptcy. The model for this study has a high explanatory power as it indicates that the combination of non-performing loans, unemployment and economic condition explain a high portion towards the variation of bankruptcy (Hilmy *et al.*, 2013).

Bankruptcy cases becomes a serious matter in Malaysia nowadays. Recently, Datuk Nancy Shukri, Minister in the Prime Minister's Department announced that were Malay race recorded the highest number in bankruptcy which is 63,011 (48 percent) followed by Chinese which is 41,732 (31 percent), Indians 17,271 (13 percent) and others 10,529 (8 percent). From that numbers, 93,036 (70.2 percent) person were male and 39,507 (29.8 percent) female. She also stated that mostly the person were age between 25 until 34 were declared as bankrupt (M. Hairulazim, 2015).

An article in Utusan Malaysia stated that, increased in the cost of living caused many of us to be depressed and not being able to manage financial wisely. Ultimately, they were unable to pay their debts and cause a person to become bankrupt. Today, more and more youth are bankrupt, because of the burden of serious debt such as credit cards, car loans and personal loans. Lifestyle beyond their often means become the main cause of the financial problems among adolescents today. According to a survey conducted by the Consumer Research and Resource Centre (CRRC) on financial behaviour and habits of young people, 37 per cent of the group admits spending more than their monthly salary. It is believed that, lack of knowledge about the financial management is the main cause of debt crisis for this group. However, in terms of the percentage of the highest bankruptcy cases for this group, it is involved that vehicle loan which is 25.21 per cent, 13.15 per cent followed by personal loans, housing loans 12.31 per cent, 11.26 per cent of business loans and 4.9 per cent for credit cards (Liaw, 2014).

## Methodology



**Figure 1 : Procedure flow of logistic regression analysis in personal loan bankruptcy.**

Figure 1 shows the flow chart on the procedure flow that we are concerning in this research study. The first thing that we are going to do is to review the factor of bankruptcy cases during 2007 until 2013. We found that personal loan problem is one of the factor that contribute to the bankruptcy cases. Then, we retrieved the data of personal loan bankruptcy cases from Malaysia Department of Insolvency (MDI) Kedah branch. We found that the variables of age, gender, race and job profession is to be consider in this study. The next step is we identify the types of data, number of samples and appropriate tests to the data. In this study, we use Likelihood ratio test and Wald test to test the significance of the unknown parameter.

For likelihood ratio test, the hypothesis testing is:

$$H_0: \beta_x = \beta_{x+1} = \beta_{p-1} = 0$$

$H_a$ : Not all of the  $\beta_x$  in  $H_0$  equal zero

The actual test statistic for the likelihood ratio test, denoted by  $G^2$  is:

$$G^2 = -2 \ln \frac{L(R)}{L(F)} = -2[\ln L(R) - \ln L(F)]$$

(1)

Where,

$L(R)$  = Likelihood function for the full model

$L(F)$  = Likelihood function for the full model

The appropriate decision rule is:

$$\text{If, } G^2 \leq \chi^2_{(1-\alpha, p-y)} \text{ accept } H_0$$

$$\text{If, } G^2 > \chi^2_{(1-\alpha, p-y)} \text{ reject } H_0$$

For Wald test, the hypothesis testing is:

$$H_0: \beta_k = 0$$

$$H_1: \beta_k \neq 0$$

An appropriate test statistic is:

$$W^2 = \left( \frac{\hat{b}_k}{s(\hat{b}_k)} \right)^2$$

Where,

$\hat{b}_k$  = estimated regression coefficient

$s(\hat{b}_k)$  = approximate estimated of standard deviation of  $\hat{b}_k$

Where  $W^2 \sim \chi^2_{(1-\alpha, 1)}$ . In accordance with this change, the decision rule must be adjusted such that

$$\text{If } W^2 \leq \chi^2_{(1-\alpha, 1)}, \text{ accept } H_0$$

$$\text{If } W^2 > \chi^2_{(1-\alpha, 1)}, \text{ reject } H_0$$

For goodness of fit test, we use Hosmer-Lemeshow test. The hypothesis are :

$$H_0: E\{Y\} = [1 + \exp(-X'\beta)]^{-1}, \text{ or the model fit is appropriate.}$$

$$H_1: E\{Y\} \neq [1 + \exp(-X'\beta)]^{-1}, \text{ or the model of fit is inappropriate.}$$

Where

$$X'\beta = \beta_0 + \beta_1 X_1 + \dots + \beta_{p-1} X_{p-1}$$

(2)

The associated test statistics is:

$$X^2 = \sum_{k=1}^K \left\{ \frac{[\sum y_{ik} - n_k \bar{\pi}_k(x_i)]^2}{n_k \bar{\pi}_k(x_i) [1 - \bar{\pi}_k(x_i)]} \right\}$$

(3)

Where  $X^2$  = chi-square random variable with  $(g-2)$  degrees of freedom. Therefore, the decision rule is:

$$\begin{aligned} \text{If } X^2 \leq \chi^2_{(1-\alpha; g-2)} & \text{ accept } H_0 \\ \text{If } X^2 > \chi^2_{(1-\alpha; g-2)} & \text{ reject } H_0 \end{aligned}$$

In order to build  $(1-\alpha)\%$  confidence interval for the slope  $\beta_k$  is

$$b_1 \pm z^* SE_{b_1}$$

(4)

A level  $(1-\alpha)\%$  confidence interval for the odds ratio  $e^{\beta_1}$  is obtained by transforming the confidence intervals for the slope:

$$(e^{b_1 - z^* SE_{b_1}}, e^{b_1 + z^* SE_{b_1}})$$

In these expressions  $z^*$  is the value for the standard normal density curve with area  $\alpha$  between  $-z^*$  and  $z^*$ .

**Results and Discussion**

**Table 1 :** Variables in the Equation of SPSS

	B	S.E.	Wald	df	Sig.	Exp(B)	95% C.I. for EXP(B)		
							Lower	Upper	
Step 1 <sup>a</sup>	Age(1)	-.107	.874	.015	1	.903	.899	.162	4.980
	Gender(1)	2.931	.567	26.685	1	.000	18.750	6.166	57.016
	Race(1)	-4.028	.667	36.460	1	.000	.018	.005	.066
	Job(1)	2.407	.953	6.373	1	.012	11.100	1.713	71.924
	Constant	1.716	1.078	2.532	1	.112	5.563		
Step 2 <sup>a</sup>	Gender(1)	2.931	.567	26.685	1	.000	18.750	6.166	57.016
	Race(1)	-3.993	.600	44.254	1	.000	.018	.006	.060
	Job(1)	2.479	.755	10.795	1	.001	11.933	2.719	52.372
	Constant	1.609	.632	6.476	1	.011	5.000		

From the output SPSS, gender ( $X_2$ ), race ( $X_3$ ) and job ( $X_4$ ) are the significant variables as shown in step 2(a). Otherwise, we also found that age ( $X_1$ ) is the variable that have been dropped from the logistic regression model. By testing the significance of the parameter by

using Wald test manually, for  $\beta_2, \beta_3$ , and  $\beta_4$ , we reject  $H_0$  and conclude  $\beta_2, \beta_3$  and  $\beta_4$  is not equal to zero.  $X_1$  has been dropped from the model due to we accept  $H_0$  and conclude that  $\beta_1$  is equal to zero. For 95% confidence interval, since the 95% confidence interval for  $\beta_1$  is including value 0, we accept  $H_0$  that the slope  $\beta_1$  is zero, while for  $\beta_2, \beta_3$ , and  $\beta_4$  we reject  $H_0$  that the slope for  $\beta_2, \beta_3$ , and  $\beta_4$  is not equal to zero.

## Conclusion

In this research, Binary Logistic Regression was explored to estimate the parameters. To assess the factors contributed to personal loans bankruptcy, we focuses on analysing the four predictors, that is age, gender, race and job profession (government or non-government sector) of bankruptcy. Multiple logistic regression is applied for this study. From the result, we obtain some conclusion for bankruptcy cases in Kedah. We found that gender, race and job profession, are the main factor contributed to the personal loan bankruptcy problem. The variable age of person does not giving any effect for personal loan bankruptcy. From the result we analysed, we can say that the odd that male to female bankruptcy is 18.750 times. Similarly, the odd that Malay bankruptcy to non-Malay bankruptcy is 0.018 times and the odd that government sectors bankruptcy to non-government sectors bankruptcy is 11.9293 times.

## References

- Dan L.Crippen(2000) *Personal Bankruptcy : Literature Review*, Congressional Budget Office Second and D Streets, S.W.xi.
- BadrulMuzammil.(2014,Feruary 20).*Debts Among Us*,The Malaysian Digest.Retrieved 15,11,2014,from <http://www.malaysiandigest.com>.
- MurshidEunos.(2014,February 26).*Bankruptcy of Young Executive*, HarianMetro,Retrieved 15,11,2014,from <http://www2.hmetro.com.my/articles/eksekutifmudabankrap/article>.
- Hairulazim Mahmud.(2015) "Orang Melayu 'juara' Muflis".Harian Metro.
- Yoon G. Lee ,Jean M. Lown, Deanna L. Sharpe (2007) *Predictors of Holding Consumer andMortgage Debt among Older Americans*, Springer Science BusinessMedia.306.
- Ning Zhu(2008) *Household Consumption and Personal Bankruptcy*,*Journal of Legal Studies*.1.
- Thomas A. Garrett (2007) *The Rise in Personal Bankruptcies: The Eighth Federal Reserve District and Beyond*, Federal Reserve Bank of St. Louis Review,17.
- HilwaHilmy, Shaliza A. Mohd Z, andNorasyikin A. Fahami (2013) *FactorsAffectingBankruptcy: The Case of Malaysia*, International Journal ofUndergraduates Studies.4-6.
- Allison Mann and Ronald Mann (2010) *Contributions of Debt and Bankruptcy toLife Course Mobility*, Columbia University and Columbia Law School.16.

Kutner, Nachtsheim, Neter and Li.(2004).*Applied Linear Statistical Models, Fourth Edition*, McGraw-Hill Irwin.551-628

Jessica Ong HaiLiaw.(2014).”Beliamuflisakibatbelanja-luarkemampuan”,Retrived 23 4,2015 from [http://ww1.utusan.com.my/utusan/Rencana/20140415/re\\_05/Belia-muflis-akibat-belanja-luarkemampuan#ixzz3aMLiA6Pw](http://ww1.utusan.com.my/utusan/Rencana/20140415/re_05/Belia-muflis-akibat-belanja-luarkemampuan#ixzz3aMLiA6Pw)

Venny Sin Woon, Chong, Jason M.S., Lam.(2012).”*An Empirical Study of Malaysian Young Adults’ Attitudes Towards Credit Card Debt: Influence of the Personal and Bank Factors*, International Conference on Management, Economics and Finance (icmef 2012) Proceeding.1-2