



UNIVERSITY OF NAIROBI

SCHOOL OF MATHEMATICS

**A MULTIVARIATE MARKOV CHAIN MODEL FOR CREDIT RISK MEASUREMENT
AND MANAGEMENT**

BY

KURUI GODFREY KIPKOECH

I56/75577/2014

**A RESEARCH PROJECT SUBMITTED IN FULFILLMENT OF THE REQUIREMENTS
FOR THE AWARD OF MASTERS OF SCIENCE DEGREE IN ACTUARIAL SCIENCE
OF THE UNIVERSITY OF NAIROBI**

November 2016

DECLARATION

I, the undersigned, declare that this is my original study and has not been submitted to any College or University for academic credit.

Information from other sources and my main respondents has been duly acknowledged.

Signed: _____ Date.....

Kipkoech Godfrey Kurui

(156/75577/2014)

This Study has been submitted for examination with our approval as University Supervisors.

Dr. Philip Ngare

Signed: _____ Date.....

DEDICATION

I dedicate this project to my Dad Philip and Mum Mary and my siblings Joan, Mercy, Keith and Dennis for love and encouragement.

ACKNOWLEDGEMENT

I thank Almighty God for giving me good health, strength and inspiration throughout my studies.

I also thank my supervisor Dr.Philip Ngare for the proper guidance and support.

I also thank my Dad Philip and Mum Mary my brothers and sisters and my friend Dr.Daniel for the encouragement and support throughout this journey.

I also thank my lecturers, classmates, friends and the entire University fraternity for the support and criticism.

Table of Contents

DECLARATION	ii
DEDICATION	iii
ACKNOWLEDGEMENT	iv
ABSTRACT.....	vii
CHAPTER ONE	1
1.1 Background	1
1.2 Problem Statement	3
1.3 Objectives	4
1.4 Justification	4
1.5 Scope	4
1.6 Limitation	4
CHAPTER TWO	5
2.1 Literature Review	5
2.1.1 Introduction	5
2.2.2 Probability of Default	5
2.2.3 Loss Given Default	5
2.2 Credit Risk Models	5
2.2.1 Structural Credit Risk Models	5
2.2.2 Reduced Form Credit Risk Models	6
CHAPTER THREE	10
3.1 METHODOLOGY	10
3.1.1 Markov chain	10
3.1.2 The period	11
3.1.3 Long Term distribution of markov chain	11
3.1.4 Irreducibility	12
3.1.5 Modeling using Markov chains	12
3.1.6 Multivariate Markov Chain Model	13
3.2 Logistic Regression	14
3.3 Cumulative Logistic Regression	15
3.4 Estimation of parameters	17
3.5 Credit Risk Measures	18
CHAPTER FOUR.....	22

4.1 Research Design	22
4.2 Data Description	22
4.3 Data Analysis	23
4.4 Illustration	24
CONCLUSION.....	31
References.....	32
APPENDIX.....	35

ABSTRACT

Logistic regression has been applied for classification and to determine the factors which affects the behavioral score of the consumer. Cumulative logistic regression with a latent variable link as link function determines the dynamic of consumers' behavioral score. A multivariate describes the dependency of credit risky assets in a portfolio. Credibility theory combines the application transition from the credit bureau with behavioral transition matrix from consumer performance and experience.

CHAPTER ONE

1.1 Background

Unregulated finance companies in Kenya are thriving despite the requirement of the Microfinance Act of 2006 that anyone conducting microfinance business has to be licensed. These ventures depend on the sanctity of contract law to preserve them in business. They are operated through ambiguous, contracts that are frequently misinterpreted by the borrower hence unrealistic interest payments. In Part II Section 9 (1) (c) of that Microfinance Act, states that a license can be revoked and the business shut down if the business being conducted is detrimental to the interests of its depositors or customers it is not clear why loan sharks in Kenya have not been challenged as the loan sharks refer to their dealings as "microfinance."

Mobile and other credit products are being developed with main objective is to enable the borrowers access small amounts of money. This also supports the borrowers to building their credit profiles and facilitate access to a higher loan limits in future. This advancement also provides a remedy for the ineffectiveness of the conventional lenders in measurement and management of credit risks.

Most of these credit providers are high a growth start-ups and therefore they must develop, test and utilize models to assess the risk within a very short time period. The main objective for this lenders is to grow their business with not only a high level of automation but also with a high degree of cautiousness and risk management. They are required to develop registered analytics that apply to their data to maintain a competitive edge in target markets and compete with established lenders. The models must be less costly and be efficient in response to loan applications thus makes credit risk modeling vital in their competitive advantage.

The credit risk models developed should be able to perform credit analysis, credit fraud identification and prevention, credit pricing, collections and portfolio management. These lenders collect and analyze their own customer-specific data and build their unique capabilities, efficiencies and competitive advantage. The process of data gathering and analysis of data can be lengthy, complex and error-prone hence the need to develop model from scratch to facilitate the analysis of proprietary data in a timely and efficient manner.

These models should be viably cost effective to diverse needs for economically actively lower income household, business and enable the lenders to make decisions in low costs. New models

that employ the increased computing power and can capture new sources of data can be developed although these new ideas are confronted with challenges with privacy laws and customers preferences.

These models provides a platform for companies to profitably serve the unbanked population and help the society toward the full inclusion goal that is these people have access to range of quality ,affordable and appropriate financial services which is a priority to most governments worldwide. The providers help population make good financial decisions, offer right noncredit products like savings and insurance and also conducting marketing and communications in effective ways that fits these populations.

Predictive analysis is vital to every lender in determination of credit worthiness of the borrowers. Historical data provided a basis for credit extension. These historical data sources like credit report and salary history aided the credit providers in assessment, predictions and risk classification by assigning scores based on the identity, ability and willingness to pay. This model utilizes new nontraditional sources of data and information to facilitate a complete understanding of the household financial needs. However these methods less effective to the unbanked populations because they have no access to formal financing hence no record of past borrowing behavior. The credit worthiness of these borrower is difficult to determine as most of these lower income have wages in cash and there is no formal savings that can be used as collateral.

Most of the population do not regular fixed payments as most are self-employed and depends on a portfolio of inconsistent income generating activities.in order for the modelers to have an effective credit scoring strategies for this group of borrowers, they are required to identify efficient sources of data and appropriate ways of accessing and converting them into them into credit insights. There are new data standards and protocols and new tools to bring together disparate data sets, matching and comparing them to bring credit risk insights.

1.2 Problem Statement

Modelling of dependency of credit risky assets is important as it has a significant impact measuring and managing credit risky assets. A model to describe this dependency should therefore be defined. Copulas and Monte Carlo techniques are the major approaches that explain the dependency of credit risky portfolios. The use of copulas was introduced by Li(2000) and its main advantages is that it can capture the dependency of credit risk when the loss distribution do not belong to the same elliptical class and it can also incorporate more than two credit risks. A multivariate markov chain model provides a not only convenient but also a natural way to describe the dependency of credit risky assets. Thomas et al (2002) suggested that a market view to be a mixture of beliefs between which is determined by both historical movements of ratings and a specified subjective view by the experts opinions.

Ching et al (2002) used a discrete homogeneous multivariate markov chain model for dependency of credit risky portfolio of securities. They applied credibility theory to combine the two sources of information from the historical data and the empirical sources and also to estimate the unknown parameters.

Our main objective is to develop a multivariate markov chain model for a portfolio of consumer loans for the unbanked populations. We demonstrate the impact of the nontraditional sources of data from the population on the generic transition matrix based on the cumulative logistic regression. We use credibility theory to combine the application and behavioral scores transitions and also estimate the unknown parameters. We derive our portfolio values from the Monte Carlo simulation and show the consistency of credit measures based on the behavioral scores.

1.3 Objectives

The objective of this study is

- i. To show impact of modelling of dependency of credit risky assets of the consumer loans in measuring and managing credit risky portfolios.
- ii. To demonstrate how the various behavioral characteristics of a borrower determines the creditworthiness of the borrower.

1.4 Justification

This research project builds a model by use of nontraditional sources of information that will enable the creditors to measure and manage the credit risk and maximize return on their investment. This model enables the unbanked population to access to a credit products despite having no credit history. This also helps the debtors build their credit profiles which enable them access credit from different sources.

1.5 Scope

This study focuses on the financial institutions in Kenya which are offering short term credit to a population which they do not have their credit history.

1.6 Limitation

Privacy issues is the key limit factor in this study.

CHAPTER TWO

2.1 Literature Review

2.1.1 Introduction

(Bonfim, 2009) and (Schmit, 2004) defined default as a situation when an debtor is unlike to pay all their obligations or past due to more than ninety days on any material credit obligation..

2.2.2 Probability of Default

(Crouhy, 2000) described credit risk modelling to be estimating probability of default (PD), the loss given default (LGD) and correlation across default and losses. Probability of default (PD) and the loss given default (LGD) are key parameters of internal rating based (IRB) approach and is central to Basel II. They enable banks to compute their capital charges for each exposure.

2.2.3 Loss Given Default

Loss given default is equivalent to the recovery rate and it is defined as the loss rate on a credit exposure if a counterparty defaults. There are three classes of loss given defaults models for individual loan or instruments namely market, workout and implied market loss given default. Market loss given default is estimated from market price of bonds or tradable loans and it is highly limited in application since after default market is only available for corporate bonds.

(Dermine, 2006) suggested the application of the workout loss given default and it is calculated from the recovered part of the exposure arising in the long running workout process and discounted to the default rate. The disadvantage for this approach is bankrupt settlements.

2.2 Credit Risk Models

There are two important models for credit risk measurement and management namely structural approach and reduced form approach which can be described as individual level reduced-form or portfolio reduced-form models

2.2.1 Structural Credit Risk Models

Merton (1974) introduced structural approach that assumes that the value of the firm assets is driven by geometric Brownian motion. The biggest disadvantage of this approach is the simplified assumptions in its derivation, but it has opened a room for extensions for instance (Geske, 1977)

Extension from the single debt maturities to various debt maturities by the compound option modelling and (Leland, 1996) allowed firms to continuously issue debts of a constant but infinite time to maturity. (Duffie, 2005) Compared (Merton, 1974) assumption that the default occurs in only at the maturity date.

Black and Cox introduced the First passage time model structural model which states that default event can happen not only at the debt maturity but also prior to that date as long as the firm asset value falls to the pre-specified barrier. With the first passage ideas, other parameters used by (Merton, 1974) were extended to be dynamic for instance (Longstaff, 1995) treated the short term risk-free interest as stochastic which converges to long term risk-free interest rate that is negatively correlated to the asset value. (Tarashev, 2005) Defined and compared the structural model as exogenous and endogenous default. The exogenous structural model defines the default as when the asset value fall below a threshold value while endogenous default, the obligor can choose time of default strategically. (Anderson, 1996) Extension allowed firms to renegotiate the term of the contract and when the default level is reached, the firm is declared as either bankrupt or a renewed contract with higher interest debt is issued.

(Shibata, 2009) Proposed a BSM structural model for banks recovery process for a company in danger of bankruptcy. This model states that in a situation where investor bankrupts, bank can either continue to run the bank or liquidate. This method defined the banks collecting process with option approach and game theory.

(Perli, 2004) Applied corporate credit risk structural model to model consumer lending. They assumed that a customer is deemed to have defaulted if his assets are below a specified threshold. This approach disregarded the key issues of the consumer defaults as consumer lending is more about cash flows and financial fraud.

(Andrade, 2007) Described structural risk model for consumer loans with the behavioral score as the proxy of credit-worthiness of the borrower. Default occurred when the value of the reputation for credit worthiness in terms of access to further credit dropped below the cost of servicing the debt.

2.2.2 Reduced Form Credit Risk Models

(Artzner, 1995) and (Duffie, D, 1996) considered the reduced- form approach assumes that default is an exogenous event and its occurrence is governed by a random point process.

Individual level reduced form models

(Altman, 1968) Proposed these models also referred to as credit scoring models and it identifies accounting variables that have statistical explanatory power to differentiate between the defaulters and non-defaulters. It applies linear or binomial models to regress the defaults and estimate the coefficients. The applicants are then scored depending on whether they are good or bad.

(Altman E. S., 1998) Studied the use of credit scoring while (Altman E. , 1997) surveyed the historical explanatory variables in credit scoring models and found that most studies used financial ratios that measures profitability, leverage and liquidity.

(Jacobson, 2003) Developed bivariate probit model and proposed a method of calculation of portfolio without bias. (Lin, 2009) Proposed a new approach by three kinds of two stage hybrid models of logistic regression artificial neural network. (Altman E. , 2005) specified a credit scoring for emerging market corporate bonds. (Luppi, 2007) applied logit model to nonprofit SMES and found that traditional accounting based credit scoring model held less explanatory power in nonprofit firms than in for profit firms.

Credit scoring models have been criticized because of its explanatory variables are based on accounting data and therefore cannot highlight the dynamic borrowers conditions. (Argawal, 2008) Study showed market based model such as structural models are better in forecasting distress than credit scoring models. It also manifested that in term of predictive accuracy using UK data, their results were almost the same as for the BSM structural model and Z-score model.

Portfolio Reduced-Form Models

(Jarrow, 1992)) introduced this models and it is associated with risk neutral technique Jarrow and Turnbull decomposed the credit risk premium and the problem of credit risk modelling became how to model probability of default (PD) and loss given default (LGD). Reduced form models can capture the firm's credit risk hence it can be specified in different stochastic process models.

Markov chain

Markov chain is a good example of portfolio reduced-form credit risk model introduced by (Jarrow R. S., 1997). It considers default as an absorbing state and that default time is a continuous markov. It also assumes a fixed probabilities for credit quality changes estimated from historical credit matrices and a fixed recovery rate (RR) in the event of default.

(Feng.D., 2008) Fitted ordered probit model to rating transition and view rating transitional probabilities as functions of latent variables as unobservable factors while (Nickell, 2000) assumed that the latent variables were derived from observable factors.

(Hurd, 2006) Generalized the markov chain to describe the dynamics of the corporate credit risk. (Kalotychou, 2006) applied ordered probit model in sovereign credit migration estimation and compared the homogeneous and heterogeneous estimators. (Gagliardini, 2005) applied ordered probit model to estimate migration correlations and suggested that the traditional cross-sectional. Monteiro et al (2006) suggested the use of finite non homogeneous continuous time semi markov process to model time dependent matrices and showed that non parametric parameters estimation of time dependent matrices.

(Kadam, 2008) extended the discrete time model by Jarrow et al (1997) continuous time markov chain in their empirical studies. Hidden markov models which is a statistical model in which the system being modelled assumed to be a markov process with an unobserved states used in forecasting of quantiles of default rates used in credit risk modelling. (Banachewicz, 2007) studied hidden markov models and tested the sensitivity of the forecasted quantiles if the underlying hidden markov models mis-specified.

(Frydman, 2008) and (Kadam, 2008) applied markov mixture model is extended to a mixture of two markov chains where the mixing is on the speed of the movement of credit ratings. The only difference is that estimation of the original markov chain was based on maximum likelihood while estimation of the mixture was based on the Bayesian estimation.

(Li, 2000) Introduced the use copula functions in modelling of dependency of credit risky securities and showed the main advantages of copula that is can be used to capture the dependency of credit risks when credit loss distribution do not belong to the same elliptical class and can incorporate the dependency of more than two credit risks. (Embrechts, 1999) Introduced the use of copula in modelling credit risks when multivariate distribution is asymmetric. (Umberto Cherubini, 2004) Discussed the use of simulation methods with various copulas for the modelling of dependent risks. (Kijima M, 2002) Applied multivariate markov chain model to replicate the development of correlated ratings of several credit risks to solve the pricing and risk measurement problems. They assumed the change in credit ratings over a period of time was driven by a single index model that consisted of the systematic and the firm specific components. The systematic component was described by a single column factor and the unknown parameters were estimated by minimizing the squared error based on the historical data only.

(Thomas. L C, 2002) Proposed that historical data alone was inadequate to describe the future movements in ratings hence an expert opinion should be incorporated. (Buhlmann,

1967) Introduced credibility theory to combine two different sources of information by determine the weights to be assign to each source. He also introduced the least square approach for estimation of credibility premiums without imposing string parameter assumptions. This theory is convenient to merge both the historical rating data and another source of information.

(Das, 2004) In their study of correlated default risk, they showed that the joint correlated default risk probabilities varied substantially over time hence a more realistic way to describe time varying behavior of risk. They also noted that the estimation procedure was complicated and less analytical.

Hu et al (2002) proposed an empirical Bayesian for estimation of transition matrices for the government to evaluate and manage risks of emerging markets .Transition matrices was estimated and assumed to be a linear combination of empirical transition matrix and model based updating matrix evaluated from an ordered probit model. They adopted empirical Bayesian techniques for estimating contingency tables and selected weights of the linear combination of prior and updating matrices by the goodness of fit chi square statistics.

Lee (1997), Bernado and Smith (2001) and Robert (2001) provide an overview and detailed discussion on the choice of the prior matrix. Transition matrix is determined based on prior knowledge of ratings of other firms with or without the same industry.

One important statistic for credit risk measurement and management is conditional expectation of aggregate loss of portfolio at future time given the available current information. Elliot et al (1997) provided the evaluation of the conditional predictive probability which is important in evaluation of credit value at risk.

Acerbi and Tasche (2001) and Hardy (1999) pointed out the addition of an adjustment term to Expected Shortfall (ES) in order to make the market coherent when loss distribution is discrete. This gave a definition of Expected Shortfall of credit portfolio at a future time given information up to current with a given probability level.

(Rosch, 2004) Took a variant of one of one factor credit metrics model and used the empirical correlation between different consumer loans and try to build the economic variables to explain differences during different parts of the business cycle. (Malik, 2010) Developed a hazard model of the time to default of consumer loans where risk factors were based on the behavioral scores, the age of the loan and economic variables.

(Belloti, 2009) Used proportional hazard model to develop a default risk model for consumer loans. Their investigation was to find out which economic variables might be most appropriate.

CHAPTER THREE

3.1 METHODOLOGY

3.1.1 Markov chain

A Markov property implies that the future value of a process is depends on the current value and is independent of the past history. A markov process is any process that satisfies the markov process. A markov chain refers to a markov process in discrete time and with a discrete state space.

Transition graph

A graphical representation of a markov chain has its states represented by circles and each arrow representing possible transitions. The probability of transition between two states can be represented in form of a transition matrix (i, j) entry with i^{th} row and j^{th} column with the probability of moving one step from state i to state j each row adding up to one at any time.

Chapman Kolmogorov

These are equations that allows calculation of general transitions probability in terms of one step probabilities $P_{i,j}^{(n,n+1)}$.

Let $P_{i,j}^{(m,n)}$ be the probability of being in state j at time n having being been in state i at time m

$$\Pr^{X_{m+1}=j / X_m=i} = P_{i,j}^{(m,m+1)}$$

The transition probability of a discrete markov chain obeys chapman-Kolmogorov equations. One step transition probability $P_{i,j}^{(n,n+1)}$ with initial probability distribution $q_k = P(X_0=k)$ is used to deduce the probability of any path.

Time –Homogeneous Markov Chain

Time homogeneous markov chain is a markov chain whose one step probabilities are time independent

$$P_{i,j}^{j,n+1} = P_{i,j}$$

The Chapman –Kolmogorov equations is given by

$$P_{i,j}^{(n-m)} = \sum_{k \in S} P_{1k}^{(i-m)} P_{kj}^{(n-1)}$$

The normalization condition $\sum_{j \in S} P_{ij} = 1$ for all i implies that each row of P must add to one $\sum_{j \in S} P_{ij} = 1$, for all i

Time inhomogeneous

Transition probabilities cannot simply be denoted P_{ij} because they will depend on the absolute values of time rather than just time difference. Transitional probabilities depend not only on the length but also when the process starts.

3.1.2 The period

A state i is said to be periodic with $d > 1$ if a return to i is possible only in a number of steps that is multiple of d . A state is aperiodic if it is not periodic. There exists $\lim_{n \rightarrow \infty} P_{ii}^{(n)}$ for the aperiodic states.

3.1.3 Long Term distribution of markov chain

A markov chain that settles in to its stationary distribution after long period of time then its distribution to tend to the invariant distribution π . If the above convergence condition holds, then $P_{ij}^{(n)}$ will be close to π_j for a larger fraction of time. We say that $\pi_j, j \in S$ is a stationary population distribution for a markov chain with transition matrix P if and if only the following condition holds

- i. $\pi_j = \sum_{i \in S} \pi_i P_{ij}$
- ii. $\pi_j \geq 0$
- iii. $\sum_{j \in S} \pi_j = 1$

This can be stated as $\pi = \pi P$ where π is a row vector.

Given a markov chain with infinite state space, a stationary distribution is found by solving $\pi = \pi P$. since this is a vector it results in a set of linear programming equations.

The above equations are said to be linearly independent implies that any four of the will always rearrange to give the remaining one redundant. By this linearity property any multiple solution of the problem above is solution and a the uniqueness of the solution comes from the normalization

$$\sum_{j \in S} \pi_j = 1$$

3.1.4 Irreducibility

A markov chain is said to be irreducible if any states can be reached from any other state. This implies that a markov chain is irreducible if given any pairs of states i, j there exist an integer n with $P_{ij}^{(n)} > 0$. An irreducible markov chain with a finite state space has a unique stationary distribution

Let $P_{ij}^{(n)}$ be n step transition probability of an irreducible, a periodic markov chain on a finite state space, then every i and j , $\lim_{n \rightarrow \infty} P_{ij}^{(n)} = \pi_j$, where π_j is stationary probability distribution.

This implies that no matter what state i you are in, the probability of ending up in state j after a very long time is the same as probability of being in state j given the stationary probability π .

A markov chain with finite state space has at least one stationary distribution. An irreducible markov chain with finite state space has a unique stationary distribution and an irreducible aperiodic markov chain with finite state space will settle down to its unique stationary distribution in the long run.

3.1.5 Modeling using Markov chains

Modeling process for the markov chain is started by fitting a stochastic model to a set of observations. Assuming that the model being fitted is time homogeneous, a state space to fit the markov model to the observed data, transitional probabilities P_{ij} is set up.

We denote x_1, x_2, \dots, x_N to be the available observations and define

- i. n_i to be number of times ($1 \leq t \leq N - 1$) such that $X_t = i$
- ii. n_{ij} to be number of times ($1 \leq t \leq N - 1$) such that $X_t = i$ and $X_{t+1} = j$

Thus n_{ij} be observed number of transitions from state i to state j and n_i be the observed number of observations from state i .

n_i allows t to go up to $N - 1$ rather than N so that it equals the number of chances of a transition out of state i and not just the number of times it is in state i .

The best estimate of $P_{ij} = \hat{P}_{ij} = \frac{n_{ij}}{n_i}$

3.1.6 Multivariate Markov Chain Model

This is given by

$$X_{t+1}^{(j)} = \sum_{k=1}^n \lambda_{jk} P^{(jk)} X_t^{(k)}, \text{ for } j = 1, 2, \dots, n.$$

The probability distribution $X_{t+1}^{(j)}$ of $Y_{t+1}^{(j)}$ given F_t depends only on $Y_t^{(1)} Y_t^{(2)} Y_t^{(3)} \dots Y_t^{(n)}$ a markov property follows that is, the conditional probability distribution of the ratings of the j^{th} credit risk at time $t + 1$ depends on the ratings of all credit risks in the portfolio at time t .

Equation above can be represent as matrix as below

$$X_{t+1} = \begin{pmatrix} X_{t+1}^{(1)} \\ X_{t+1}^{(2)} \\ \vdots \\ X_{t+1}^{(n)} \end{pmatrix} = \begin{pmatrix} \lambda_{11}P^{(11)} \lambda_{12}P^{(12)} & \dots & \lambda_{1n}P^{(1n)} \\ \lambda_{21}P^{(21)} \lambda_{22}P^{(22)} & \dots & \lambda_{2n}P^{(2n)} \\ \vdots & \ddots & \vdots \\ \lambda_{n1}P^{(n1)} \lambda_{n2}P^{(n2)} & \dots & \lambda_{nn}P^{(nn)} \end{pmatrix} \times \begin{pmatrix} X_t^{(1)} \\ X_t^{(2)} \\ \vdots \\ X_t^{(n)} \end{pmatrix} \dots$$

$$X_{t+1} = QX_t$$

PROPOSITION 1

Suppose that $P^{(jk)}$ ($1 \leq j, k \leq n$) are irreducible and $\lambda_{jk} > 0$ then there is a vector

$X = [X^{(1)}, X^{(2)}, \dots, X^{(n)}]^T$ Such that $X_{t+1} = QX_t$ and $\sum_{i=1}^n [X^{(j)}]^i = 1, (1 \leq j, k \leq n)$ where $[\cdot]^i$ is corresponding vector.

Vector X is has a stationary probability distribution of the ratings of all credit risks in the portfolio i.e, for each j in $X^{(j)}$ represents probability of ratings in the long run.

3.2 Logistic Regression

Logistic regression describes the relationship with a binary response with one or more explanatory variables by applying the logit transformation to the dependent variable.

A simple logistic model is defined as the logit (Y) = natural logarithm of odds ratio given as below;

$$Y = \ln \left[\frac{\pi(x)}{1-\pi(x)} \right] = \alpha + \beta X.,$$

The antilogarithm of the above gives the probability of an occurrence of outcomes of interest as follows.

$$\pi = \frac{e^{\alpha+\beta X}}{1 + e^{\alpha+\beta X}}$$

π is the probability of outcome of interest

α is the intercept

β is regression coefficient

The *logit of Y and X* in the above equation is linear while the probability of *Y and X* is nonlinear hence to make the relationship of *Y and X* linear, the above equation is transformed by taking the natural logarithm.

The sign of the regression coefficient β determines the direction of relationship between *logit of Y and X*.

The simple logistic regression above can be extended to multiple predictors variables X_1, X_2, \dots, X_p , then logistic regression of *Y* is as above

$$Y = \ln \left[\frac{\pi(x)}{1 - \pi(x)} \right] = \alpha + X_1\beta_1 + X_2\beta_2 + \dots + X_p\beta_p$$

Then probability of ($Y = \text{outcome of Interest} / X_1 = x_1, X_2 = x_2$) =

$$\pi = \frac{e^{\alpha+X_1\beta_1+X_2\beta_2+\dots+X_p\beta_p}}{1 + e^{\alpha+X_1\beta_1+X_2\beta_2+\dots+X_p\beta_p}}$$

α and β Are estimated by the Maximum Likelihood method. This method maximizes the likelihood of reproducing the data given the parameters estimates.

A null hypothesis underlying this model states that all β_s are all zeros. A rejection of this null hypothesis implies that at least one of the β_s doesn't equal zero in the population. This means logistic regression equation predicts the outcome better than the mean of the outcome response *Y*. Interpretation of results is rendered using the odds ratio for both categorical and continuous predictors.

3.3 Cumulative Logistic Regression

This is a common method for analysis univariate ordered categorical data and it estimates the effects of the explanatory variables on the log odds of selecting lower response than the higher response categories.

$$\text{If } p(Y \leq j/x) = \pi_1(x) + \pi_2(x) + \dots + \pi_j(x)$$

Then

Cumulative logistic regression logit is defined as

$$\begin{aligned} \text{logit } p(Y \leq j/x) &= \log \left(\frac{p(Y \leq j/x)}{(1 - p(Y \leq j/x))} \right) \\ &= \log \left(\frac{\pi_1(x) + \pi_2(x) + \dots + \pi_j(x)}{\pi_{j+1}(x) + \pi_{j+2}(x) + \dots + \pi_J(x)} \right) \end{aligned}$$

Each cumulative logit uses all j responses categories. A model that uses logit $p(Y \leq j)$ is alone an ordinary logit model for binary response in which categories 1 to j form one outcome and category $j + 1$ to J form the second.

A model that simultaneously uses all cumulative logits is

$$\text{logit } p(Y \leq j/x) = \alpha + \beta X, j = 1 \dots J - 1$$

Each cumulative logit has its own intercept (α) and it is increasing in j since $p(Y \leq j/x)$ increases in j for fixed x . The logit is an increasing function of this probability.

Cumulative model can be expressed as a latent variable of the form

$$\begin{aligned} y_i^* &= \sum_{k=1}^K \beta_k X_{ik} + \varepsilon_i \\ \text{And } y_i &= \begin{cases} 1 & \text{if } y_i^* \leq \alpha_1 \\ 2 & \text{if } \alpha_2 \leq y_i^* \leq \alpha_1 \\ 3 & \text{if } \alpha_3 \leq y_i^* \leq \alpha_2 \\ \vdots & \vdots \\ J-1 & \text{if } \alpha_{j-2} \leq y_i^* \leq \alpha_{j-1} \\ J & \text{if } \alpha_{j-1} \leq y_i^* \leq \alpha_j \end{cases} \end{aligned}$$

If the distribution for error term ε_i is the logistic distribution then, the model is of the form

$$\ln \left[\frac{\Pr(Y_i \leq b)}{\Pr(Y_i > b)} \right] = \alpha_b - \sum_{k=1}^K \beta_k X_{ik} \quad b = j \dots \dots j - 1$$

- Y_i is the response of the i^{th} individual with y_i observed variable and y_i^* latent variable)
- X_{ik} - the $k - th$ explanatory variable for the i^{th} individual
- J is the number of ordered categories of the independent variable.
- α'_b s Partitions intercept (cut points) indicates the logarithm of odds of selecting lower rather than higher categories when all explanatory variables are set to zero.
- $\alpha = (\alpha_1 \dots \dots \alpha_{j-1})$ is the vector of parameters in which $\alpha_1 \leq \alpha_2 \leq \dots \leq \alpha_{j-1}$
- $\beta = (\beta_1 \dots \dots \beta_k)$ - the vector of regression coefficients for the explanatory variable?
- K - is the number explanatory variable.

If y^* is a latent variable that underlies a regression model for a continuous variable and has cdf $G(y^* - n)$ where y^* vary around parameter $n = \text{mean}$ that depends on X such that $n(x) = \beta(x)$. Suppose that $-\infty \leq \alpha_1 \leq \alpha_2 \leq \dots \leq \alpha_{j-1} \leq \infty$ are cutpoints of the continuous scale such that the observed response Y satisfies $Y = j$ if $\alpha_{j-1} \leq Y_i^* \leq \alpha_j$ that is Y fall in the category j when latent variable fall in the j^{th} interval of values.

$$p(Y \leq j/x) = p(Y^* \leq j/x) = G(\alpha_j - \beta \cdot X)$$

Then this implies that the link is the inverse of the cdf for Y^* applies to $p(Y \leq j/x)$.

If $y^* = \beta X + \epsilon$ where G is the logistic regression the inverse of G is the logit link that gives the proportional odds.

The linear predictor $\sum_{k=1}^K \beta_k X_{ik}$ is subtracted from the intercepts and the positive coefficient indicates an increased probability of higher response category. Cumulative logistic model assumes that the effect of different explanatory variables are fixed across all $(j - 1)$ Partitions of the ordinal response.

Proportional odds model has the same effect β for each logit and for a continuous fixed predictor x and a fixed j , the response curve is a logistic regression curve with binary response with outcomes $\leq j$ and $Y > j$. The response curves have the same rate of increase or decrease but are horizontally displaced. The same parameters B occurs for the effects of Y regardless of how cut points chop up the scale for the latent variable and this makes possible to compare estimates from the studies using different response variables.

For $j < k$ the curve for $p(Y \leq k)$ is a curve for $p(Y \leq j)$ translated by $\frac{(\alpha_k - \alpha_j)}{\beta}$ units in the X direction then

$$p\left(\left(Y \leq k / X = x\right) = p\left(\left.Y \leq j\right/ X = x + \frac{(\alpha_k - \alpha_j)}{\beta}\right)\right)$$

If cumulative logit model given by

$$\text{logit } p(Y \leq j/x) = \alpha + \beta X, j = 1 \dots \dots \dots J - 1$$

Then

$$\text{logit} \left(p(Y \leq j/x_1) - p(Y \leq j/x_2) \right) = \beta(x_1 - x_2)$$

3.4 Estimation of parameters

A common transition matrix for credit risks assets assumed to be based on the prior information the behavioral score is used.

The estimate of $N_e^{(jk)}$ of $N^{(jk)}$ is given by $N_e^{(jk)} = W_{jk}M^{(jk)} + (1 - W_{jk})N^{(jk)}$ for $j, k = 1, 2 \dots \dots n$ where

$$0 \leq j, k \leq 1, \text{ for } j, k = 1 \leq , 1, 2 \dots \dots n$$

$Q^{(jk)}$ is the prior matrix for estimation of $N^{(jk)}$

The proposition above it states that multivariate markov chain has a stationary distribution X . Then vector X be taking the proportion of occurrence of each state in categorical time series of ratings.

Denote the estimate of X by $\check{X} = (\check{X}^{(1)} \check{X}^{(2)} \dots \dots \check{X}^{(n)})^T$

Note from the proposition above Q matrix is given by

$$\begin{pmatrix} \lambda_{11}P_e^{(11)} & \lambda_{12}P_e^{(12)} & \dots & \dots & \dots & \lambda_{1n}P_e^{(1n)} \\ \lambda_{21}P_e^{(21)} & \lambda_{22}P_e^{(22)} & \dots & \dots & \dots & \lambda_{2n}P_e^{(2n)} \\ & & \vdots & & & \\ & & & \vdots & & \\ \lambda_{n1}P_e^{(n1)} & \lambda_{n2}P_e^{(n2)} & \dots & \dots & \dots & \lambda_{nn}P_e^{(nn)} \end{pmatrix} \hat{X} \approx \hat{X}$$

Let $\lambda_{jk}^1 = \lambda_{jk} W_{jk}$ and $\lambda_{jk}^2 = \lambda_{jk}(1 - W_{jk})$. it can be checked that $\lambda_{jk}^1 + \lambda_{jk}^2 = \lambda_{jk}$, for each $j, k=1, 2, \dots, n$

$$\begin{aligned} & \min_{\lambda^1, \lambda^2} O_j \\ & \left\{ \text{subject to} \begin{pmatrix} O_j \\ O_j \\ O_j \\ O_j \\ \cdot \\ O_j \\ O_j \end{pmatrix} \geq \widehat{X}^{(j)} - B_j \begin{pmatrix} \lambda_{j1}^1 \\ \lambda_{j1}^2 \\ \lambda_{j2}^1 \\ \lambda_{j2}^2 \\ \cdot \\ \lambda_{jn}^1 \\ \lambda_{jn}^2 \end{pmatrix}, \begin{pmatrix} O_j \\ O_j \\ O_j \\ \cdot \\ \cdot \\ O_j \\ O_j \end{pmatrix} \geq -\widehat{X}^{(j)} + B_j \begin{pmatrix} \lambda_{j1}^1 \\ \lambda_{j1}^2 \\ \lambda_{j2}^1 \\ \lambda_{j2}^2 \\ \cdot \\ \lambda_{jn}^1 \\ \lambda_{jn}^2 \end{pmatrix} \right. \end{aligned}$$

$$O_j \geq 0$$

$\sum_{k=1}^n (\lambda_{jk}^1 + \lambda_{jk}^2) = 1, \lambda_{jk}^1 \geq 0$ and $\lambda_{jk}^2 \geq 0, \forall j, k$ where

$$B_j = Q^{(j1)} \widehat{X}^{(1)} P^{(j1)} / Q^{(j2)} \widehat{X}^{(2)} / P^{(j2)} \widehat{X}^{(2)} / \dots / Q^{(jn)} \widehat{X}^{(n)} / P^{(jn)} \widehat{X}^{(n)}$$

3.5 Credit Risk Measures

Define

$$\begin{aligned} E_p(\mathcal{L}_{t+1}(\mathbf{Y}_{t+1}) / \mathcal{F}_t) &= \sum_{j=1}^n \sum_{i=1}^m \mathbf{E}_P(\langle L_{t+1}^j, \mathbf{e}_i \rangle I\{\boldsymbol{\omega} \in \Omega \mid \mathbf{Y}_{t+1}^{(j)}(\boldsymbol{\omega}) = \mathbf{e}_i\} / \mathcal{F}_t) \\ &= \sum_{j=1}^n \sum_{i=1}^m \langle L_{t+1}^j, \mathbf{e}_i \rangle \rho(\{Y_{t+1}^{(j)} = e_i\} / \mathcal{F}_t) \end{aligned}$$

$$P_{t+1/t}^{(j)} := P_{t+1/t}^{(j1)}, P_{t+1/t}^{(j2)}, \dots, P_{t+1/t}^{(jn)}$$

$$P_{t+1/t}^{ji} := P(\{Y_{t+1}^{(j)} = e_i\} / \mathcal{F}_t) = E_p(\langle Y_{t+1}^{(j)}, \mathbf{e}_i \rangle / \mathcal{F}_t)$$

$$= E_p(\langle Y_{t+1}^{(j)}, \mathbf{e}_i \rangle) |_{Y_t = (e_{i_1}, e_{i_2}, \dots, e_{i_n})}$$

From equation (1) below,

$$X_{t+1}^{(j)} = \sum_{k=1}^n \lambda_{jk} P^{(jk)} X_t^{(k)}, \text{ for } j = 1, 2, \dots, n.$$

The unknown parameters in the above equation can be estimated as above:

$$X_{t+1}^{(j)} = \sum_{k=1}^n \lambda_{jk} P^{(jk)} X^{(k)} \approx \sum_{k=1}^n (\lambda_{jk}^{\sim 1} Q^{(jk)} + \lambda_{jk}^{\sim 2} \hat{P}^{(jk)}) X_t^{(k)}, \text{ for } j = 1, 2, \dots, n.$$

Let $[V]^i$ denote the i^{th} element of the column vector V . then for each $j = 1, 2, \dots, n$ and $i = 1, 2, \dots, n$. We have,

$$\begin{aligned} P_{t+1/t}^{(ji)} &= E_{\rho}(\langle Y_{t+1}^{(j)}, e_i \rangle) |_{Y_t=(e_{i_1}, e_{i_2}, \dots, e_{i_n})} \\ &= \rho(\{Y_{t+1}^{(j)} = e_i\}) |_{Y_t=e_{i_1}, e_{i_2}, \dots, e_{i_n}} \\ &= [X_{t+1}^{(j)}]^i |_{Y_t=(e_{i_1}, e_{i_2}, \dots, e_{i_n})} = [X_{t+1}^{(j)}]^i |_{X_t=(e_{i_1}, e_{i_2}, \dots, e_{i_n})} \\ &= \left[\sum_{k=1}^n \lambda_{jk} P^{(jk)} X^{(k)} \right]^i |_{X_t=e_{i_1}, e_{i_2}, \dots, e_{i_n}} \approx \left[\sum_{k=1}^n (\lambda_{jk}^{\sim 1} Q^{(jk)} + \lambda_{jk}^{\sim 2} \hat{P}^{(jk)}) X_t^{(k)} \right]^i |_{X_t=(e_{i_1}, e_{i_2}, \dots, e_{i_n})} \end{aligned}$$

This also implies that

$$\begin{aligned} E_{\rho}(\mathcal{L}_{t+1}(Y_{t+1})/\mathcal{F}_t) &= \sum_{j=1}^n \sum_{i=1}^m \langle L_{t+1}^j, e_i \rangle \rho(\{Y_{t+1}^{(j)} = e_i\}/\mathcal{F}_t) \\ &= \sum_{j=1}^n \sum_{i=1}^m \langle L_{t+1}^j, e_i \rangle P_{t+1/t}^{(ji)} \approx \sum_{j=1}^n \sum_{i=1}^m \langle L_{t+1}^j, e_i \rangle \left[\sum_{k=1}^n (\lambda_{jk}^{\sim 1} Q^{(jk)} + \lambda_{jk}^{\sim 2} \hat{P}^{(jk)}) X_t^{(k)} \right]^i |_{X_t=(e_{i_1}, e_{i_2}, \dots, e_{i_n})} \end{aligned}$$

The joint conditional predictive distribution of Y_{t+1} given the information set \mathcal{F}_t is key in evaluation of credit value at risk (VaR).

$$P_{t+1/t}^{(j)} := (P_{t+1/t}^{j1}, P_{t+1/t}^{j2}, \dots, P_{t+1/t}^{jn})^T \approx \sum_{k=1}^n (\lambda_{jk}^{\sim 1} Q^{(jk)} + \lambda_{jk}^{\sim 2} \hat{P}^{(jk)}) X_t^{(k)} |_{X_t=(e_{i_1}, e_{i_2}, \dots, e_{i_n})}$$

$Y_{t+1}^{(1)}, Y_{t+1}^{(2)}, \dots, Y_{t+1}^{(n)}$ are conditionally independent given \mathcal{F}_t or Y_t thus the joint conditional predictive distribution $P_{t+1/t}$ of Y_{t+1} , given the information \mathcal{F}_t can be completely determine by

$P_{t+1/t} := (P_{t+1/t}^{j1}, P_{t+1/t}^{j2}, \dots, P_{t+1/t}^{jn})^T$, where $P_{t+1/t}$ is $(n \times m)$ - dimensional probability matrix.

Then conditional predictive probability that the aggregate loss \mathcal{L}_{t+1} equals $\mathcal{L}_{t+1}(\bar{k})$ is given by:

$$\begin{aligned}
& \rho(\mathcal{L}_{t+1} = \mathcal{L}_{t+1}(\bar{k}) / \mathcal{F}_t) \\
&= \sum_{(i_1, i_2, \dots, i_n) \in I_{t+1, \bar{k}}} \prod_{j=1}^n P_{t+1/t}^{(j i_j)} \\
&\approx \sum_{(i_1, i_2, \dots, i_n) \in I_{t+1, \bar{k}}} \left\{ \prod_{j=1}^n \left[\sum_{k=1}^n (\lambda_{jk}^{\sim 1} Q^{(jk)} + \lambda_{jk}^{\sim 2} \hat{P}^{(jk)}) X_t^{(k)} \Big|_{X_t = (e_{i_1}, e_{i_2}, \dots, e_{i_n})} \right]^{i_j} \right\}
\end{aligned}$$

The VaR of the portfolio with probability level $\alpha \in (0,1)$ at time $t+1$ given the market information \mathcal{F}_t is given by $VaR_{\alpha, \rho}(\mathcal{L}_{t+1} / \mathcal{F}_t) := \inf\{L \in \mathcal{R} / \rho(\mathcal{L}_{t+1} \geq L / \mathcal{F}_t) \leq \alpha\}$

Let K^* denote a positive integer in $\{1, 2, \dots, M\}$ such that

$$\begin{aligned}
\rho(\mathcal{L}_{t+1} \geq \mathcal{L}_{t+1}(K^*) / \mathcal{F}_t) &= \sum_{k=K^*}^M \rho(\mathcal{L}_{t+1} = \mathcal{L}_{t+1}(\bar{k}) / \mathcal{F}_t) \leq \alpha \\
\rho(\quad) &= \sum_{k=K^*+1}^M \rho(\mathcal{L}_{t+1} = \mathcal{L}_{t+1}(\bar{k}) / \mathcal{F}_t) > \alpha
\end{aligned}$$

Then we have

$$VaR_{\alpha, \rho}(\mathcal{L}_{t+1} / \mathcal{F}_t) = \mathcal{L}_{t+1}(K^*)$$

$$ES_{\alpha}(\mathcal{L}_{t+1} / \mathcal{F}_t) = \frac{1}{\alpha} E_p(\mathcal{L}_{t+1} \mathbf{I}_{\{\mathcal{L}_{t+1} \geq \mathcal{L}_{t+1}(K^*)\}}) + A(\alpha)$$

Where adjustment term $A(\alpha)$ is given by

$$A(\alpha) := \mathcal{L}_{t+1}(K^*) \left[1 - \frac{\rho(\mathcal{L}_{t+1} \geq \mathcal{L}_{t+1}(K^*) / \mathcal{F}_t)}{\alpha} \right]$$

$$\begin{aligned}
ES_{\alpha}(\mathcal{L}_{t+1} / \mathcal{F}_t) &= \frac{1}{\alpha} \left[\sum_{k=K^*}^M \mathcal{L}_{t+1}(\bar{k}) \rho(\mathcal{L}_{t+1} \geq \mathcal{L}_{t+1}(\bar{k}) / \mathcal{F}_t) \right. \\
&\quad \left. - \mathcal{L}_{t+1}(K^*) * \rho(\mathcal{L}_{t+1} \geq \mathcal{L}_{t+1}(K^*) / \mathcal{F}_t) - \alpha \right]
\end{aligned}$$

$$= \frac{1}{\alpha} \left\{ \sum_{k=K^*}^M \mathcal{L}_{t+1}(k) \left(\sum_{(i_1, i_2, \dots, i_n) \in I_{t+1, \check{k}}} \prod_{j=1}^n P_{t+1/t}^{j i_j} \right) - \mathcal{L}_{t+1}(K^*) \right. \\ \left. * \left[\sum_{k=K^*}^M \left(\sum_{(i_1, i_2, \dots, i_n) \in I_{t+1, \check{k}}} \prod_{j=1}^n P_{t+1/t}^{j i_j} \right) - \alpha \right] \right\}$$

$$\approx \frac{1}{\alpha} \left\{ \sum_{k=K^*}^M \mathcal{L}_{t+1}(k) \left\{ \sum_{(i_1, i_2, \dots, i_n) \in I_{t+1, \check{k}}} \left[\sum_{k=1}^n (\lambda_{jk}^{-1} Q^{(jk)} + \lambda_{jk}^{-2} \hat{P}^{(jk)}) X_t^{(k)} \Big|_{X_t=(e_{i_1}, e_{i_2}, \dots, e_{i_n})} \right]^{i_j} \right\} - \right. \\ \left. \mathcal{L}_{t+1}(K^*) \left\{ \sum_{k=K^*}^M \left\{ \sum_{(i_1, i_2, \dots, i_n) \in I_{t+1, \check{k}}} \left[\sum_{k=1}^n (\lambda_{jk}^{-1} Q^{(jk)} + \lambda_{jk}^{-2} \hat{P}^{(jk)}) X_t^{(k)} \Big|_{X_t=(e_{i_1}, e_{i_2}, \dots, e_{i_n})} \right]^{i_j} \right\} - \right. \right. \\ \left. \left. \alpha \right\} \right\}$$

CHAPTER FOUR

4.1 Research Design

The research design used was the descriptive design which involves obtaining information about a current status of a situation to describe “what Exists” with respect to variables in a situation.

4.2 Data Description

A data from a loan issuing firm with 30,000 customers with their demographic features and their behavioral characteristics in terms of repayment status, loan limit and the amount paid in three months was used. A logistic regression was performed on a sample of 220 using the R program as follows

Results

The results to the logistic regression is a per the attached in the Appendix 2.a summary of is as shown on the table below;

Coefficients:

Coefficients	Estimate	p-value
intercept	-0.949	0.343
sex	-0.544	0.587
education	0.586	0.558
marriage	-0.39	0.697
Payment status	-4.590	0.444
Outstbal	-1.172	0.241
Bill	1.115	0.265
Age	0.839	0.401

4.3 Data Analysis

On test of significance we compare the p value from Wald test and conclude that for any p value less than .5 as significant while for greater than .5 is removed and the test is carried out again. The value of gender, education status and marriage status predictors are all eliminated.

All the factors that remained in the second regression are all significant and shows how each relate with response at hand variable.

4.4 Illustration.

The marginal probability of cumulative logistic regression is given by

$$p(y_i = j) = \begin{cases} F(\delta_1) & \text{if } j = 1 \\ F(\delta_1) - F(\delta_1) & \text{if } j = 1 \text{ nad } j - 1 \\ F(\delta_1) \leq J \end{cases}$$

Assuming that the estimated behavioral transition matrix for this consumers is given as below

$$N = \begin{bmatrix} 0.60 & 0.28 & 0.09 & 0.03 \\ 0.07 & 0.10 & 0.30 & 0.53 \\ 0.371 & 0.465 & 0.015 & 0.014 \\ 0.278 & 0.038 & 0.302 & 0.382 \end{bmatrix},$$

And the estimated empirical transition for the matrices as above,

$$\begin{aligned} \dot{M}^{(11)} &= \begin{bmatrix} \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ 0 & \frac{7}{9} & \frac{1}{8} & 0 \\ 0 & \frac{2}{9} & \frac{7}{8} & 0 \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} \end{bmatrix}, \dot{M}^{(12)} = \begin{bmatrix} \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ 0 & \frac{6}{11} & \frac{1}{6} & 0 \\ 0 & \frac{5}{11} & \frac{5}{6} & 0 \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} \end{bmatrix}, \dot{M}^{(21)} = \begin{bmatrix} \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ 0 & \frac{8}{9} & \frac{1}{4} & 0 \\ 0 & \frac{1}{9} & \frac{3}{4} & 0 \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} \end{bmatrix}, \\ \dot{M}^{(22)} &= \begin{bmatrix} \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ 0 & \frac{2}{11} & \frac{1}{6} & 0 \\ 0 & \frac{9}{11} & \frac{5}{6} & 0 \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} \end{bmatrix}, \end{aligned}$$

The state probability for the following states in the long run is as follows

$$X^{(1)} = \begin{bmatrix} 0 \\ 0.5 \\ 0.5 \\ 0 \end{bmatrix}, \quad X^{(2)} = \begin{bmatrix} 0 \\ 0.61 \\ 0.39 \\ 0 \end{bmatrix}$$

$$X_{t+1}^{(j)} = \sum_{k=1}^2 (\tilde{\lambda}_{jk}^1 \hat{N}^{(jk)} + \tilde{\lambda}_{jk}^2 M^{(jk)}) X^{(k)}$$

$$X_{t+1}^{(1)} = \sum_{k=1}^2 (\tilde{\lambda}_{1k}^1 \hat{N}^{(1k)} + \tilde{\lambda}_{1k}^2 M^{(1k)}) X^{(k)}$$

$$= (\tilde{\lambda}_{11}^1 \hat{N}^{(11)} + \tilde{\lambda}_{11}^2 M^{(11)}) X^{(1)} + (\tilde{\lambda}_{12}^1 \hat{N}^{(12)} + \tilde{\lambda}_{12}^2 M^{(12)}) X^{(2)}$$

$$= X_{t+1}^{(1)} = \tilde{\lambda}_{11}^1 \hat{N}^{(11)} X^{(1)} + \tilde{\lambda}_{11}^2 M^{(11)} X^{(1)} + \tilde{\lambda}_{12}^1 \hat{N}^{(12)} X^{(2)} + \tilde{\lambda}_{12}^2 M^{(12)} X^{(2)}$$

We formulate our estimation problem as follows

$$\min_{\tilde{\lambda}_{jk}^1, \tilde{\lambda}_{jk}^2} \max_i [\tilde{\lambda}_{11}^1 \hat{N}^{(11)} X^{(1)} + \tilde{\lambda}_{11}^2 M^{(11)} X^{(1)} + \tilde{\lambda}_{12}^1 \hat{N}^{(12)} X^{(2)} + \tilde{\lambda}_{12}^2 M^{(12)} X^{(2)} - X^{(1)}]^i$$

$$\text{Subject to: } \tilde{\lambda}_{11}^1 + \tilde{\lambda}_{11}^2 + \tilde{\lambda}_{12}^1 + \tilde{\lambda}_{12}^2 = 1$$

$$\tilde{\lambda}_{11}^1, \tilde{\lambda}_{11}^2, \tilde{\lambda}_{12}^1, \tilde{\lambda}_{12}^2 \geq 0$$

The above equation can be reformulated as below;

$$\text{Let } O_j = \max_i \left[\tilde{\lambda}_{11}^1 \begin{bmatrix} 0.5 & 0 & 0 & 0.5 \\ 0 & 0.78 & 0.125 & 0 \\ 0 & 0.22 & 0.875 & 0 \\ 0.5 & 0 & 0 & 0.5 \end{bmatrix} \begin{bmatrix} 0 \\ 0.5 \\ 0.5 \\ 0 \end{bmatrix} + \tilde{\lambda}_{12}^1 \begin{bmatrix} 0.60 & 0.28 & 0.09 & 0.03 \\ 0.07 & 0.10 & 0.30 & 0.53 \\ 0.371 & 0.465 & 0.015 & 0.014 \\ 0.278 & 0.038 & 0.302 & 0.382 \end{bmatrix} \begin{bmatrix} 0 \\ 0.5 \\ 0.39 \\ 0 \end{bmatrix} \right. \\ \left. + \tilde{\lambda}_{12}^2 \begin{bmatrix} 0.5 & 0 & 0 & 0.5 \\ 0 & 0.55 & 0.33 & 0 \\ 0 & 0.45 & 0.67 & 0 \\ 0.5 & 0 & 0 & 0.5 \end{bmatrix} \begin{bmatrix} 0 \\ 0.61 \\ 0.39 \\ 0 \end{bmatrix} + \tilde{\lambda}_{12}^2 \begin{bmatrix} 0.60 & 0.28 & 0.09 & 0.03 \\ 0.07 & 0.10 & 0.30 & 0.53 \\ 0.371 & 0.465 & 0.015 & 0.014 \\ 0.278 & 0.038 & 0.302 & 0.382 \end{bmatrix} \begin{bmatrix} 0 \\ 0.61 \\ 0.39 \\ 0 \end{bmatrix} \right]$$

$$\min_{\tilde{\lambda}_{jk}^1, \tilde{\lambda}_{jk}^2} O_j$$

Subject to:

$$O_j \geq (X^{(j)}) - B_j \begin{pmatrix} \tilde{\lambda}_{11}^1 \\ \tilde{\lambda}_{11}^2 \\ \tilde{\lambda}_{12}^1 \\ \tilde{\lambda}_{12}^2 \end{pmatrix}, O_j \geq (-X^{(j)}) + B_j \begin{pmatrix} \tilde{\lambda}_{11}^1 \\ \tilde{\lambda}_{11}^2 \\ \tilde{\lambda}_{12}^1 \\ \tilde{\lambda}_{12}^2 \end{pmatrix}$$

$$O_j \geq 0$$

$$\sum_{k=1}^n (\tilde{\lambda}_{jk}^1 + \tilde{\lambda}_{jk}^2) = 1 \quad \tilde{\lambda}_{jk}^1, \tilde{\lambda}_{jk}^2 \geq 0 \quad \forall j, k$$

B_j is given by $Q^{(j1)} \hat{X}^{(1)} P^{(j1)} / Q^{(j2)} \hat{X}^{(2)} / P^{(j2)} \hat{X}^{(2)} / \dots / Q^{(jn)} \hat{X}^{(n)} / P^{(jn)} \hat{X}^{(n)}$

For the example below we have $B_j = Q^{(11)} \hat{X}^{(1)} P^{(11)} \hat{X}^{(1)} / Q^{(12)} \hat{X}^{(2)} / P^{(12)} \hat{X}^{(2)} / Q^{(22)} \hat{X}^{(2)} / P^{(22)} \hat{X}^{(2)}$

Given below

$$M^{(11)} X^{(1)} = \begin{bmatrix} 0.5 & 0 & 0 & 0.5 \\ 0 & 0.78 & 0.125 & 0 \\ 0 & 0.22 & 0.875 & 0 \\ 0.5 & 0 & 0 & 0.5 \end{bmatrix} * \begin{bmatrix} 0 \\ 0.5 \\ 0.5 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0.451389 \\ 0.548611 \\ 0 \end{bmatrix}$$

$$N^{(11)} X^{(1)} = \begin{bmatrix} 0.60 & 0.07 & 0.371 & 0.278 \\ 0.28 & 0.10 & 0.465 & 0.038 \\ 0.09 & 0.30 & 0.15 & 0.302 \\ .03 & 0.53 & 0.014 & 0.382 \end{bmatrix} * \begin{bmatrix} 0 \\ 0.5 \\ 0.5 \\ 0 \end{bmatrix} = \begin{bmatrix} 0.4505 \\ 0.3825 \\ 0.125 \\ 0.042 \end{bmatrix}$$

$$M^{(12)} X^{(2)} = \begin{bmatrix} 0.5 & 0 & 0 & 0.5 \\ 0 & 0.55 & 0.33 & 0 \\ 0 & 0.45 & 0.67 & 0 \\ 0.5 & 0 & 0 & 0.5 \end{bmatrix} * \begin{bmatrix} 0 \\ 0.61 \\ 0.39 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0.564815 \\ 0.435185 \\ 0 \end{bmatrix}$$

$$N^{(12)} X^{(2)} = \begin{bmatrix} 0.60 & 0.07 & 0.371 & 0.278 \\ 0.28 & 0.10 & 0.465 & 0.038 \\ 0.09 & 0.30 & 0.15 & 0.302 \\ 0.03 & 0.53 & 0.014 & 0.382 \end{bmatrix} * \begin{bmatrix} 0 \\ 0.61 \\ 0.39 \\ 0 \end{bmatrix} = \begin{bmatrix} 0.468167 \\ 0.364167 \\ 0.119444 \\ 0.048222 \end{bmatrix}$$

This reformulation problem is therefore as shown below,

$$\min_{\tilde{\lambda}_{jk}^1, \tilde{\lambda}_{jk}^2} O_j$$

Subject to

$$(O_1) \geq 0 - 0. \tilde{\lambda}_{11}^1 - 0.4505 \tilde{\lambda}_{11}^2 - 0. \tilde{\lambda}_{12}^1 - 0.468167 \tilde{\lambda}_{12}^2$$

$$(O_1) \geq \frac{1}{2} - 0.451389 \tilde{\lambda}_{11}^1 - 0.3825 \tilde{\lambda}_{11}^2 - 0.564815 \tilde{\lambda}_{12}^1 - 0.364167 \tilde{\lambda}_{12}^2$$

$$(O_1) \geq \frac{1}{2} - 0.548611 \tilde{\lambda}_{11}^1 - 0.125 \tilde{\lambda}_{11}^2 - 0.435185 \tilde{\lambda}_{12}^1 - 0.119444 \tilde{\lambda}_{12}^2$$

$$(O_1) \geq 0 - 0.0 \tilde{\lambda}_{11}^1 - 0.042 \tilde{\lambda}_{11}^2 - 0.0 \tilde{\lambda}_{12}^1 - 0.042822 \tilde{\lambda}_{12}^2$$

$$1 = \tilde{\lambda}_{11}^1 + \tilde{\lambda}_{11}^2 + \tilde{\lambda}_{12}^1 + \tilde{\lambda}_{12}^2$$

$$\tilde{\lambda}_{11}^1, \tilde{\lambda}_{11}^2, \tilde{\lambda}_{12}^1, \tilde{\lambda}_{12}^2 \geq 0$$

The solution this optimization problem can be the obtained by the excel solver as follows

$$X_{t+1}^{(1)} = 0.09985M^{(11)}X^{(1)} + 0.90015N^{(12)}X^{(2)}$$

$$X_{t+1}^{(2)} = N^{(12)}X^{(2)}$$

Computation of Credit Value at Risk and Estimated shortfall

The model developed above is used to generate the predictive probability distributions for evaluating credit risk measures.

$$\text{Let } L_{t+1}(Y_{t+1}^{(1)}, Y_{t+1}^{(2)}) = L_{t+1}(Y_{t+1}^{(1)}) + L_{t+1}(Y_{t+1}^{(2)})$$

For each of $j = 1, 2$ above, the rating class $j^{(th)}$ and $Y_{t+1}^{(j)}$ at time $t + 1$ can take values in set of unit basis vector $(e_1 e_2 \dots e_4) \in R^4$.

Considering a unit vector $[0, 1]$ and its uniform partition $U_{i=1}^8 p_i$ where $p_i = \left(\frac{i-1}{4}, \frac{i}{4}\right)$

Assuming that for each $j = 1,2$ and $i = 1,2 \dots 4$ the loss from the j^{th} asset $Y_{t+1}^{(j)}(Y_{t+1}^{(i)})$ given $Y_{t+1}^{(j)} = e_i$ therefore takes the interval p_i for each $i = 1,2 \dots 4$

This implies that $Y_{t+1}^{(j)} = e_i$ the loss from the j^{th} asset $Y_{t+1}^{(j)}(Y_{t+1}^{(i)})$ at time $t + 1$ can take values in $p_1 = \left[0, \frac{1}{4}\right]$. The simulated results is as below

1	0.1235	0.2672	0.3949	0.3221
2	0.0289	0.1857	0.3453	0.3773

The ordered aggregate losses from the credit portfolio at time $t + 1$ based on the the simulated values below are given above;

0.753,0.851,0.942,1.003,1.043,1.435,1.236,1.347,1.548,1.679,1.777,1.803,1.900,1.998,2.038,2.591,2.65.

The predictive probability from the equation derived above is as below

.0001,0.0053,0.0032,0.300,0.035,0,0,0,0.046,0.0014,0.091,0.48,0.021,0.003,0.002,0.012

Credit Risk Measures

To evaluate the credit value at risk we choose a value K^* such that it satisfies this two equations

$$(\mathcal{L}_{t+1} \geq / \mathcal{L}_{t+1}(K^*) / \mathcal{F}_t) = \sum_{k=K^*}^M \rho(\mathcal{L}_{t+1} = \mathcal{L}_{t+1}(\bar{k}) / \mathcal{F}_t) \leq \alpha$$

$$\rho(\mathcal{L}_{t+1} \geq / \mathcal{L}_{t+1}(K^*) + 1 / \mathcal{F}_t) = \sum_{k=K^*+1}^M \rho(\mathcal{L}_{t+1} = \mathcal{L}_{t+1}(\bar{k}) / \mathcal{F}_t) > \alpha$$

For our illustration the two equations are satisfied at the point $K^* = 12$ and thus the

$$VaR_{\alpha, \rho}(\mathcal{L}_{t+1} / \mathcal{F}_t) = \mathcal{L}_{t+1}(K^*) = 1.9$$

We can estimate the expected shortfall as follows;

$$ES_{\alpha}(\mathcal{L}_{t+1} / \mathcal{F}_t) = \frac{1}{\alpha} E_p(\mathcal{L}_{t+1} \mathbf{I}_{\{\mathcal{L}_{t+1} \geq / \mathcal{L}_{t+1}(K^*)\}}) + A(\alpha)$$

Where adjustment term $A(\alpha)$ is given by

$$A(\alpha) := \mathcal{L}_{t+1}(K^*) \left[1 - \frac{\rho(\mathcal{L}_{t+1} \geq \mathcal{L}_{t+1}(K^*)/\mathcal{F}_t)}{\alpha} \right]$$

$$\begin{aligned} ES_\alpha(\mathcal{L}_{t+1}/\mathcal{F}_t) &= \frac{1}{\alpha} \left[\sum_{k=K^*}^M \mathcal{L}_{t+1}(\bar{k}) \rho(\mathcal{L}_{t+1} \geq \mathcal{L}_{t+1}(\bar{k})/\mathcal{F}_t) - \mathcal{L}_{t+1}(K^*) * \rho(\mathcal{L}_{t+1} \geq \mathcal{L}_{t+1}(K^*)/\mathcal{F}_t) - \alpha \right] \\ &= \frac{1}{\alpha} \left\{ \sum_{k=K^*}^M \mathcal{L}_{t+1}(k) \left(\sum_{(i_1, i_2, \dots, i_n) \in I_{t+1, \bar{k}}} \prod_{j=1}^n P_{t+1/t}^{j i_j} \right) - \mathcal{L}_{t+1}(K^*) \right. \\ &\quad \left. * \left[\sum_{k=K^*}^M \left(\sum_{(i_1, i_2, \dots, i_n) \in I_{t+1, \bar{k}}} \prod_{j=1}^n P_{t+1/t}^{j i_j} \right) - \alpha \right] \right\} \end{aligned}$$

$$\begin{aligned} &\approx \frac{1}{\alpha} \left\{ \sum_{k=K^*}^M \mathcal{L}_{t+1}(k) \left\{ \sum_{(i_1, i_2, \dots, i_n) \in I_{t+1, \bar{k}}} \left[\sum_{k=1}^n (\lambda_{jk}^{-1} Q^{(jk)} + \lambda_{jk}^{-2} \hat{P}^{(jk)}) X_t^{(k)} \Big|_{X_t=(e_{i_1}, e_{i_2}, \dots, e_{i_n})} \right]^{i_j} \right\} \right\} - \\ &\mathcal{L}_{t+1}(K^*) \left\{ \sum_{k=K^*}^M \left\{ \sum_{(i_1, i_2, \dots, i_n) \in I_{t+1, \bar{k}}} \left[\sum_{k=1}^n (\lambda_{jk}^{-1} Q^{(jk)} + \lambda_{jk}^{-2} \hat{P}^{(jk)}) X_t^{(k)} \Big|_{X_t=(e_{i_1}, e_{i_2}, \dots, e_{i_n})} \right]^{i_j} \right\} \right\} - \\ &\alpha \left. \right\} \end{aligned}$$

Using $\alpha = 0.05$, and value $K^* = 12$, then

$$= \frac{1}{0.05} \left[\sum_{k=12}^{16} \mathcal{L}_{t+1}(K^*) \{p(\mathcal{L}_{t+1}(K^*)/\mathcal{F}_t)\} \right] =$$

$$\sum_{k=12}^{16} \mathcal{L}_{t+1}(K^*) \{p(\mathcal{L}_{t+1}(K^*)/\mathcal{F}_t)\} = 0.948992$$

$$= \frac{1}{.05} * 0.948992 = 18.9784$$

$$\frac{1}{0.05} \left[\mathcal{L}_{t+1}(K^*) \left\{ \sum_{k=12}^{16} \{p(\mathcal{L}_{t+1}(K^*)/\mathcal{F}_t)\} - \alpha \right\} \right]$$

$$\mathcal{L}_{t+1}(K^*) = 1.900$$

$$\begin{aligned} & \frac{1}{0.05} \left[\mathcal{L}_{t+1}(K^*) \left\{ \sum_{k=12}^{16} \{p(\mathcal{L}_{t+1}(K^*)/\mathcal{F}_t)\} - \alpha \right\} \right] \\ &= \frac{1}{0.05} \{1.900(0.518 - 0.05)\} \\ &= 17.784 \end{aligned}$$

$$\begin{aligned} ES_{\alpha}(\mathcal{L}_{t+1}/\mathcal{F}_t) &= 18.9784 - 17.784 \\ &= 1.19584 \end{aligned}$$

When $\alpha = 0.01$ then

$$K^* = 14$$

$$VaR_{\alpha, \rho}(\mathcal{L}_{t+1}/\mathcal{F}_t) = \mathcal{L}_{t+1}(K^*) = 2.048$$

$$\begin{aligned} & \frac{1}{0.01} \left[\sum_{k=14}^{16} \mathcal{L}_{t+1}(K^*) \{p(\mathcal{L}_{t+1}(K^*)/\mathcal{F}_t)\} \right] = 3.6992 \\ & \frac{1}{0.01} \left[\mathcal{L}_{t+1}(K^*) \left\{ \sum_{k=14}^{16} \{p(\mathcal{L}_{t+1}(K^*)/\mathcal{F}_t)\} - \alpha \right\} \right] = 0.7000 \end{aligned}$$

$$\begin{aligned} ES_{\alpha}(\mathcal{L}_{t+1}/\mathcal{F}_t) &= 3.6992 - 0.700 \\ &= 2.9992 \end{aligned}$$

CONCLUSION

Since both value at risk (Var) and expected shortfall (ES) increases as the probability level decreases we conclude that the both are consistent with their understanding.

Cumulative link model classifies observations into categories with unknown distance between them. It also provides a regression framework like the linear models but it treats the outcome response as categorical. Cumulative logits combines many ordinary logistic regression into single model and it utilizes the available information. Logistic regression has been used here to define the prior based on the credit bureau and the behavioral score of the consumer. The dynamics of credit scores is described by the multivariate markov chain and its transition probabilities are defined by the cumulative link distribution.

This study notes that cumulative logistic regression utilizes the available information and further suggest the application of risk neutral technique in the specification of the transition matrix.

References

- Altman, E. (1968). Financial ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy. *Journal of Finance*, 589-609.
- Altman, E. (1997). *Business Failure Classification Models: An International Survey*. New York: Wiley.
- Altman, E. (2005). An emerging market credit scoring system for corporate bonds: Emerging Markets Review. 311-323.
- Altman, E. S. (1998). Credit risk management : Developments over the last 20 years. *Journal of banking and Finance*, 1721-1742.
- Anderson, R. Y. (1996). Corporate Bond Yield Spreads and Term Structure.
- Andrade, F. (2007). Structural models in consumer credit. *European Journal of Operational Research Society*, 1569-1581.
- Argawal, V. T. (2008). Comparing the performance of the market-based and accounting -based bankruptcy prediction models. *Journal of Banking and Finance* , 1541-1551.
- Artzner, P. e. (1995). Default Risk Premium and Incomplete markets. *Journal of Mathematical Finance*, 187-195.
- Banachewicz, K. (2007). Quintile Forecasting for Credit Risk Management Using Possibly Misspecified Hidden Markov Models. *Tinbergen Institute Discussion Papers*.
- Belloti, T. (2009). Credit scoring with Macroeconomic Variables using Survival Analysis. *Journal of Operation Research Society*, 1699-1707.
- Bonfim, D. (2009). Credit Risk Drivers: Evaluating the contribution of firm level information and microeconomics dynamics. *Journal of Banking and Finance*, 281-299.
- Buhlmann, H. (1967). Experience Rating and Credibility Theory. 199-207.
- Crouhy, M. G. (2000). A comparative Analysis of Current Credit Risk Models. *Journal of Banking and Finance*, 59-117.
- Das, S. ., (2004). Correlated Default Risk, . *Working Paper*.
- Dermine, j. C. (2006). Bank Loan Losses-given Default: A Case Study. *Journal of banking and Finance*, 1219-1243.
- Duffie, D. (2005). Credit Risk Modelling with affine Processes. *Journal of Banking and Finance*, 2751-2802.
- Duffie, D, S. M. (1996). Recursive Valuation of Defaulted Securities and the Timing of Resolution of Uncertainty . *Tha Annals of Applied Probability*, 1075-1090.

- Embrechts, p. M. (1999). Correlation and Dependence in risk management:properties and pitfalls. *Risks*, May ,69-71.
- Feng.D., G. (2008). The ordered qualitative model for credit ratings transitions. *Journal Empirical Finance*, 111-130.
- Frydman, H. (2008). Credit ratings dynamics and Markov mixture models. *Journal of Banking and Finance*, 1062-1075.
- Gagliardini, P. (2005). Migration correlation:Definition and efficient estimation. *Journal of banking and Finance*, 865-894.
- Geske, R. (1977). The Evaluation of Corporate Liabilities as Compound Options. *Journal Financial Economics*, 63-81.
- Hurd, T. K. (2006). Affine Markov chain model of multiform credit migration . *Journal of Credit Risk*, 3-29.
- Jacobson, T. (2003). Bank lending policy,credit scoring and value at risk. *Journal of Banking and Finance*, 615-633.
- Jarrow, R. (1992). Credit Risk: Drawing the Analogy,. *Risk Magazine*.
- Jarrow, R. S. (1997). A markov model for the term structureof credit risk spreads. *Review of Financial Studies*, 481-523.
- Kadam, A. (2008). Bayesian inference for issuer heterogeneity in credit ratings migration. *Journal of Banking and Finance*, 2267-2274.
- Kalotychou, E. (2006). On sovereign Credit Migration:A study of Alternative Estimators and RatingDynamics. *Society of Computational Economics*.
- Kijima M, K. K. (2002). A Multivariate Markov Model for Simulating Correlated Defaults. *Journal of Risk*, 1-32.
- Leland, H. T. (1996). Optimal Capital Structure, Endogeneous Bankruptcy , and Term Structure of Credit Risk of Credit Spreads . *Journal of Finance*, 987-819.
- Li, D. (2000). On default correlation. *On a copula function approach Journal of Fixed Income*, 43-54.
- Lin, S. (2009). A new two-stagehybridapproach of credit risk in banking industry,Expert systems with Applications.
- Longstaff, F. (1995). A simple Approach to Valuing Risky Fixed and Floating Rate Debt. *Journal of Finance*, 789-819.

- Luppi, B. M. (2007). Credit risk and Base I: Are non-profit firms financially different? *Working papers Series*.
- Malik, M. (2010). Modelling of the credit risk of portfolio of consumer loans. *Journal of Operation Research Society*, 411-420.
- Merton, R. (1974). On the Pricing of the Corporate Debt: The Risk Structure of Interest Rates. *Journal of Finance*, 449-470.
- Nickell, P. S. (2000). Stability of rating transitions. *Journal of Banking and Finance*, 203-227.
- Perli, R. (2004). Economic and regulatory capital allocation for revolving retail exposures. *Journal of Banking and Finance*, 789-809.
- Rosch, D. (2004). Forecasting Retail Portfolio Credit Risk . *Journal of Risk Finance*, 16-32.
- Schmit, M. (2004). Credit Risk in Leasing Industry. *Journal of Banking and Finance*, 811-833.
- Shibata, T. T. (2009). Dynamic Model of Credit Risk in Relationship Lending: A Game -Theoretical Real Options Approach.
- Tarashev, N. (2005). An Empirical Evaluation of Structural Credit Risk Models . *Bank of International Settlements*, 1-53.
- Thomas, L C, A. D. (2002). A hidden Markov Chain Model for the Term Structure of Credit Risk Spreads. *International Review of Finance*, 311-329.
- Umberto Cherubini, E. L. (2004). *Copulas Methods in Finance*. New York: John Wiley & Sons .

APPENDIX

Appendix I

The results was as follows for the first logistic regression

> Summary (glm1)

Call:

Glm (formula = default ~ sex + education + marriage + payment status +
Bill + outstbal + age, family = binomial ("logit"), data = data excel)

Min	1Q	Median	3Q	Max
-2.22377	-0.68723	-0.52287	-0.05687	2.49167

Coefficients:

	Estimate	Std. Error	z value	Pr (> z)
(Intercept)	-1.428e+00	1.505e+00	-0.949	0.343
Sex	-1.977e-01	3.637e-01	-0.544	0.587
Education	1.600e-01	2.728e-01	0.586	0.558
Marriage	-1.508e-01	3.871e-01	-0.390	0.697
Payment status	7.666e-01	1.670e-01	4.590	4.44e-06 ***
Bill	-6.450e-05	5.504e-05	-1.172	0.241
Outstbal	6.333e-05	5.680e-05	1.115	0.265
Age	2.085e-01	2.485e-01	0.839	0.401

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 240.07 on 218 degrees of freedom

Residual deviance: 199.56 on 211 degrees of freedom

(2 observations deleted due to missingness)

AIC: 215.56

Number of Fisher Scoring iterations: 6

Second Logistic regression

Summary (glm3)

Call:

Glm (formula = default ~ payment status + bill + outstbal + age,
Family = binomial ("logit"), data = data excel)

Deviance Residuals:

Min	1Q	Median	3Q	Max
-2.18855	-0.67412	-0.54488	-0.05045	2.46704

Coefficients:

	Estimate	Std. Error	z value	Pr (> z)
(Intercept)	-1.869e+00	7.116e-01	-2.626	0.00865 **
Payment status	7.867e-01	1.648e-01	4.774	1.81e-06 ***
Bill	-6.929e-05	5.531e-05	-1.253	0.21034
Outstbal	6.851e-05	5.705e-05	1.201	0.22983
Age	2.652e-01	2.238e-01	1.185	0.23590

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 240.07 on 218 degrees of freedom

Residual deviance: 200.41 on 214 degrees of freedom

(2 observations deleted due to messiness)

AIC: 210.41

Number of Fisher Scoring iterations: 6

Data Description`				Relative Frequencies in percentage	
Variable	Description	Categories	Score	Defaulters	Non Defaulters
Id	ID for each customer				
Default	Default Status in the next month	1. Default 2. Non default	1 0		
Sex	Gender	1. Male 2. Female	1 0		
Education	Education level	1. Graduate school 2. University 3. High school 4. Others	1 2 3 4		
Age	Age groups	1. ≤ 25 2. $26 \leq x \leq 29$ 3. $30 \leq x \leq 39$ 4. $40 \leq x \leq 49$ 5. $50 \leq x \leq 59$ 6. > 60	1 2 3 4 5 6		

Marriage	Marriage status	1. Married 2. Single 3. Others	1 2 3		
payment status	Payment status as at previous month	1. Paid duly 2. payment delay>1month 3. payment delayed >2months 4. payment delayed>3months	1 2 3 4		
Loan Bal	Outstanding Loan Balance as at previous month	1. <=0 2. =0 3. =>0	1 2 3		
Loan limit	Loan limit as at the previous month	1. 10000<=y<=100000 2. 100000<y<=200000 3. 200000<=y<=350000 4. >350000	1 2 3 4		