MODELLING CREDIT RISK FOR PERSONAL LOANS USING KAPLAN-MEIER TECHNIQUE

ABSTRACT

Kaplan-Meier technique was adopted to estimate time to default on loan obligations concerning personal loans for male and female applicants. For each group, an account was observed for a period of 36 months. The life of an account is measured from the month it was opened until the account becomes ‘bad’ or it is closed or until the end of observation. The account is considered bad if payment is not made for two consecutive months in line with the industry practice. If the account does not miss two payments and is closed or survives beyond the observation period, it is considered to be censored. This project aims to estimate not only the probability of defaulting but also when a borrower is likely to default. This assists the lender to fairly price risks and focus on ultimate profitability.

Log-rank test will be used to compare time to default for male and female applicants to come up with a common hypothesis.

Table of Contents

CHAPTER ONE

INTRODUCTION

1.1. Background of The Study

Over the past decade, a number of the world's largest banks have developed sophisticated systems in an attempt to model the credit risk arising from important aspects of their business lines. Such models are intended to aid banks in quantifying, aggregating and managing risk across geographical and product lines. The outputs of their models also play important roles in bank's risk management and performance measurement process, including performance-based compensation, customer profitability analysis, risk-based pricing, active portfolio management and capital structure decisions.

Credit risk is the possibility of a loss resulting from a borrower's failure to repay a loan or meet contractual obligations. Traditionally, it means that a lender may not receive the owed principal and interest, which results in an interruption of cash flows and increased cost of collection. There are many different forms of credit. Common examples include car loans, mortgages, personal loans and lines of credit. Essentially, when the bank or other financial institution makes a loan, its "credits" money to the borrower, who must pay it back at a future date.

Lending dates back to at least ancient Mesopotamia when agricultural communities would borrow seeds and animals with the promise to repay once the crops were harvested or the animals gave birth. In modern society, lending occurs whenever someone swipes a credit card to buy a cup of coffee, takes out a mortgage to buy a home, or uses student loans to attend a university. Common lenders include financial institutions, that build a business model around lending money. The borrower pays a price for taking out the loan in form of interest. If the lender feels there's a higher risk of not being paid back by a borrower, like with a new start-up business, they will charge that borrower a higher interest rate. Lower-risk borrowers pay lower interest rates.

In retail banking, credit risk models assist in the decision of whether to grant credit to an application or not. Traditional credit risk models aim at determining a customer's probability of defaulting on loan repayment. This has been changing over the recent years towards choosing the customers of highest profit. This means that it now becomes important to know not only if but when a customer will default *Thomas, L. (2002)*.

Traditionally, credit scoring aimed at distinguishing good payers from bad payers at the time of application. The timing when customers default is also important to investigate since it can provide the bank with the ability to do profit scoring. Analyzing when customers default is typically tackled using survival analysis techniques. The major strength of survival analysis is that it allows censored data to be incorporated into the model. This translates in the consumer credit context as a customer who never defaults, pays off the loan early so an event of interest is not observed. In credit risk modelling, the event of interest is default, thus modelling time to default on loan obligations.

Modelling of credit risk using survival analysis was first introduced by *Narain. (1992)*. *Thomas, L. (2002)* further developed the model. *Narain. (1992)* applied the survival model on 24 months of loan data. The result showed that the survival analysis approach provides more detailed and relevant information for credit management than the conventional approaches. *Narain. (1992)* applied the technique by using the accelerated life exponential model to 24 months loan data. The author showed the model estimated the number of failures at each failure time well. Then a scorecard was built using multiple regression, and it was shown to be superior to conventional credit scoring methods in that a better credit granting decision could be made if the score was supported by the estimated survival times. The research by *Thomas, L. (2002)* also did a comparison of exponential, Weibul and Cox non-parametric models with logistic regression and found that survival analysis techniques are competitive with, and sometimes superior to, traditional logistic regression approach. Furthermore, the idea of competing risks was employed when two possible outcomes were considered: default and early settlement. *Orgler (1970)* applied regression analysis in a model for commercial loans. *Wiginton (1980)* was one of the first to publish credit scoring results using logistic regression. It was compared with discriminant analysis. *Leonard (1993)* also applied logistic regression in evaluating commercial loans. *Durand (1941)* pioneered the use of discriminant analysis for credit scoring. Decision tree and rule was adopted by *Makowski (1985)* and *Mehta (1968)* for credit scoring. Other techniques include K-Nearest Neighbor Classifiers which was used by *Chatterjee and Barcun (1970*) and *Henley and Hand (1996)*. *Baesens. (2003)* studied the use of Bayesian network classifiers to rate borrowers. Linear programming was applied by *Hardy and Adrian (1985)* and compared it with other statistical approaches.

1.2. Problem statement

Over the past years, the Kenyan financial market has experienced growing liquidity, which has caused banks to rigorously market various loan products. This has given rise to the need to review the banks' credit granting criteria to reflect the growing volume of loan portfolio and to respond to the current global credit crunch. However, research on credit risk has surprisingly received insignificant attention from both practitioners and scholars in Kenya and the larger African continent. Over the years, banks have perpetually used traditional credit scoring techniques to rate loan applicants.

There is a system used by lenders to gauge the creditworthiness of potential borrowers known as "5 C's" (*Troy Segal ,2022*). The 5 C's are important because lenders use these factors to determine whether to approve you for a financial product Lenders also use them to set your loan rates and loan terms. The 5 Cs are character, capacity, capital, collateral and conditions. This system weighs five characteristics of the borrower and conditions of the loan, attempting to estimate the chance of default and consequently the risk of a financial loss to the lender. This method incorporates both qualitative and quantitative measures. Each lender has its own method for analyzing a borrower's credit worthiness. Most lenders use the 5 C's when analyzing individual or business credit applications. The problem arises when an applicant is categorized as a "bad" which means that there is high risk of default, most times they are not granted the loan. By incorporating survival techniques with this system, it is possible to estimate the probability of defaulting and also the time the default is likely to happen. This may help the lender to develop a suitable offer that better suits a specific applicant.

A number of studies have been carried out on the issue of credit risk modelling using different approaches. A limited number of studies have applied survival analysis techniques but none has used Kaplan-Meier method to analyse credit risk. This research intended to model probability of servicing loans and hazard rates for both male and female borrowers using this method. Furthermore, existing credit scoring models classify borrowers into different risk categories but cannot provide any information on when the borrower is likely to default. It is more informative for the lender not only to know the probability of defaulting but also when the default is likely to happen. This helps to fairly price risks and improve the focus on ultimate profitability.

For instance, if the lender knows that a group of loan applicants are bad type, instead of rejecting their applications, it may grant loans to them at a higher interest rate, as long as the term of the loan is shorter than the likely time to default. Thus some “bad” applicants can also be considered as profitable propositions. The traditional structural models currently used by most institutions are unable to capture this as illustrated above. It is also worth noting that the banks have traditionally and consistently categorized borrowers in terms of some risk groups. Accordingly, there is an apparent need to test whether these classifications constitute homogenous risk groups.

1.3. Justification

Survival analysis is a relatively new application that offers an advantage of predicting time to the event of interest. Therefore, it lays the foundation for estimating the applicants’ profitability. This is superior to the traditional logistic approach which assumes that accounts that do not experience default are “good” (non-defaulting accounts), while survival analysis treats such accounts in a more comprehensive manner, as those that proved to be “good” so far.

The outcome of the research is expected to put to light the reliability and consistency or otherwise, of the survival methods of data analysis. Likewise, credit risk analysts may draw from the research on better approaches to classifying loan applicants based on risk characteristics. Besides, the research outcome may also serve as the basis for setting risk premium to be loaded on to the base rates.

Survival analysis is also more preferred as it has the following strengths:

1. Survival analysis is able to account for censoring, unlike other techniques.
2. Unlike linear regression, survival analysis has a binary outcome, which is more realistic.
3. It analyses time to default rather than mere probability of defaulting.

1.4. Purpose of the study

To illustrate that by using survival analysis techniques such as the Kaplan-Meier model over the traditional techniques to model credit risk, we get a better result by lenders offering better and marketable risks that suits each type of borrower and eventually improving the focus on ultimate profitability.

1.5. Objectives

1.5.1. General objective

The broad objective of this research project was to use Kaplan-Meier survival model to generate default probabilities at various points in time. The study also intended to perform a test of equality of the two risk groups, namely male and female applicants.

1.5.2. Specific objectives

The specific objectives include:

1. To estimate time to default using Kaplan-Meier estimator for each risk group.

2. To determine hazard rate for each risk group on the basis of Kaplan-Meier estimator.

3. To test the statistical significance of the differences in the survival curves for each risk group based on log-rank tests.

1.5.3. Research Question

The questions that are to be answered at the end of this project is;

1. Does Kaplan-Meier technique produce better results in terms of evaluation of credit risk?
2. Is there any significant difference in the survival curves for male and female applicants?
3. Does gender affect credit risk?

CHAPTER TWO
LITERATURE REVIEW

INTRODUCTION

In this chapter, we are going to talk about survival analysis, Kaplan Meier model (Product Limit Approach), censoring, survival analysis curves and log rank test. We will also review the most important works about failure prediction methodologies. After analyzing the most popular alternative statistical techniques that can be used to develop credit risk models, we will also focus on works that have investigated the problem of modelling credit risk for personal loans using other survival techniques.

In the literature, various techniques have been described. Most of them are Classification Models in which the main purpose is to distinguish good type borrowers from the bad type. Another technique is application of the survival model. Its major advantage is to predict the time-to-default, which is never known under the Classification Models. Since survival analysis is the proposed modelling, it will be deeply investigated and later, brief reviews of Classification Models. The review ranges from historical background 2.5 to Survival Modelling 2.16.

2.1 Survival Analysis Technique

Survival analysis is a branch of statistics for analyzing the expected duration of time until one or more events happen for example loan default, death, failure of machines. It attempts to answer certain questions such as; what is the proportion of the population that will survive past a certain time? Of those that will survive, what rate will they die or default? Can multiple causes of death or default be taken into account? How do particular circumstances or characteristics increase or decrease the probability of surviving?

More generally, survival analysis involves the modelling of time to event data; in this context, loan default is considered as an event in survival analysis.

2.2 Kaplan-Meier Model

Kaplan Meier Model, also known as the Product-Limit Approach, is a non-parametric technique of estimating and plotting the survival probability as a function of time. It is often the first step in carrying out the survival analysis as it is the simplest approach and requires least assumptions. (*Mills*, *2011*). In 1985, Edward L. Kaplan and Paul Meier collaborated to publish a seminal paper on how to deal with incomplete observations. Subsequently, Kaplan Meier curves have become a familiar way of dealing with differing survival times (time to event), especially when not all the subjects continue in the study. “Survival” times need to relate to actual survival with loan default being the event of interest (*Stalpers & Kaplan, 2018*).

First assumption, censoring is unrelated to the outcome. The Kaplan Meier method assumes that the probability of censoring is unrelated to the outcome of interest. Circumstances that affect the study (positively or negatively) are not assumed to change the baseline, and the third is that the events occurred at the time specified. However, sometimes we don’t know the exact date of an event but only its status at each observation. (*Tsai, 1999*).

In credit risk, we never know the true survival functions. That is why we use the Kaplan Meier estimator to approximate the true survival function from the collected data. The estimator is defined as the fraction of observations who survived in certain amount of time under the same circumstances. However, in preparing the Kaplan Meier survival analysis, each subject is characterized by three variables: their observation time, their status at the end of observation time (event occurrence or censored), the study group they are in.

These components may be displaced in a table. For the construction of survival times and probabilities and curves, the serial times for individual subjects are arranged from the shortest to the longest, without regard to when they entered the study. (*Wang, 2010*).

Some of the common mistakes that can occur while working with the Kaplan Meier estimator includes removing censored data, interpreting the ends of the curve and accounting for one predictor. It might be tempting to remove censored data as it can significantly alter the shape of the Kaplan Meier curve. However, this can lead to severe biases so, we should always include it while fitting the model. It is important to pay special attention when interpreting the end of the survival curves, as any big drops close to the end of the study can be explained by only a few observations reaching this point of time. This should also be indicated by wider confidence intervals.

The Kaplan Meier estimator is a univariable method as it approximates the survival function using at most one variable/predictor. The results can be easily biased that is either exaggerating or missing the signal that is caused by the so-called omitted-variable bias, which then causes the analysis to assume that the potential effects or multiple predictors should be attributed only to the single one, which we take into account.

Kaplan Meier estimate adopts some conventions. First, the hazard of experiencing the event is zero at all durations except those that occurs in the sample under study. Secondly, the hazard experiencing the event at any particular duration is given by; the number of individuals experiencing the event at that duration divided by the risk set at that duration. Finally, persons that are censored are removed from the observation at the duration at which censoring takes place.

2.3 Censoring

In survival analysis, censored observations contribute to the total number at risk up to the time they ceased to be followed. (*Prinja, S., Gupta, N., & Verma, R., 2010*).

The main reasons for censoring arise from several situations. When some Individuals not experiencing the event when the study is over, some lost to follow-up during the study period and others withdraws from the study.

There are several types of censoring: random censoring, right censoring (most common type of censoring), occurs when a subject's known event of interest occurs sometime after the recorded follow-up period. Right censoring is a bit unrealistic since it's rare to have all subjects enrolled at the same time in a study. A subject is left censored if it is known that the event of interest occurs some time before the recorded follow-up period. Generally, left censoring occurs when a person's true survival time is less than or equal to the observed survival time. A subject is interval censored if it is known that the event of interest occurs between two times but the exact time of failure is not known. An important assumption is made to make appropriate use of censored data. We assume that censoring is independent or unrelated to the likelihood of developing the event of interest.

In Kaplan Meier curves, the length of the horizontal line along the x-axis of serial times represents the survival duration for that interval. The vertical distances between the horizontals are important because they illustrate the change in cumulative probabilities as the curve advances. The interpretation of survival curve is quite simple, the y-axis represents the probability that the subject still has not experienced the event of interest after surviving up to time t, represented on the x-axis. The non-continuous nature of the Kaplan Meier curves emphasizes that they are not smooth functions but rather stepwise estimates; thus, calculating a point survival is difficult. (*Rich et al, 2010*).

To add, some of the advantages of Kaplan Meier are; It gives the average view of the population, only the information about the time to event is required meaning we can use any categorical features describing groups and also, it automatically handles class imbalance, as virtually any proportion of default to censored events is acceptable. The limitations aligned to this model is that we cannot simultaneously account for multiple factors for observation and the assumption are of independence between censoring and survival can be inapplicable or unrealistic. In conclusion, even with a few disadvantages the Kaplan Meier survival curves are a great place to start off while conducting survival analysis. (*Mills, 2011*).

2.4 Log-Rank Test

Log rank test is a statistical test to test the null hypothesis of no difference in survival between two or more independent groups. The test compares the entire survival experience between groups and can be thought of as a test whether the survival curves are identical or not. Survival curves are estimated for each group using the Kaplan Meier method and compared statistically using the log rank test. (*Peto & Pike, 1973*).

There are several variations of log rank test but in our case, we will use linked close to Chi-square statistics test of independence that compares the observed to expected numbers of events numbers of events at each timepoint over the follow-up period.

2.5 History of Credit Scoring

Credit scoring is essentially a way of recognizing the different groups in a population when one cannot see the characteristic that separates the groups but only related ones. This idea of discriminating between groups in a population was introduced in statistics by Fisher (*Fisher 1936*). He sought to differentiate between two varieties of iris by measurements of the physical size of the plants and to differentiate the origins of skulls using their physical measurements. David Durand (*Durand 1941*) in 1941 was the first to recognize that one could use the same techniques to discriminate between good and bad loans.

His was a research project for the US National Bureau of Economic Research and was not used for any predictive purpose. At the same time some of the finance houses and mail order firms were having difficulties with their credit management. Decisions on whether to give loans or send merchandise had been made judgmentally by credit analysts for many years. However, these credit analysts were being drafted into military service and there was a severe shortage of people with this expertise. So, the firms got the analysts to write down the rules of thumb they used to decide to whom to give loans (*Johnson 1992*). These rules were then used by non-experts to help make credit decisions - one of the first examples of expert systems. It did not take long after the war ended for some folk to connect these two events and to see the benefit of statistically derived models in lending decisions. The first consultancy was formed in San Francisco by Bill Fair and Earl Isaac in the early 1950s and their clients at that time were mainly finance houses retailers and mail order firms

The arrival of credit cards in the late 1960s made the banks and other credit card issuers realize the usefulness of credit scoring. The number of people applying for credit cards each day made it impossible both in economic and manpower terms to do anything but automate the lending decision. When these organizations used credit scoring, they found that it also was a much better predictor than any judgmental scheme and default rates would drop by 50% or more - see (*Myers 1963*) for an early report on such success or Churchill et al. (*Churchill, Nevin, Watson 1977*) for one from a decade later.

The only opposition came from those like Capon (*Noel Capon 1982*) who argued '‘that the brute force empiricism of credit scoring offends against the traditions of our society”. He felt that there should be more dependence on credit history and it should be possible to explain why certain characteristics are needed in a scoring system and others are not. The event that ensured the complete acceptance of credit scoring was the passing of the Equal Credit Opportunity Acts (*ECOA 1975, ECOA 1976*) in the US in 1975 and 1976. These outlawed discriminating in the granting of credit unless the discrimination could be statistically justified. It is not often that lawmakers provide long term employment for anyone but lawyers but this ensured that credit scoring analysis was to be a growth profession for the next 25 years. This has proved to be the case and still is the case. So, the number of analysts in the UK has doubled even in the last four years. In the 1980s the success of credit scoring in credit cards meant that banks started using scoring for their other products like personal loans, while in the last few years scoring has been used for home loans and small business loans.

Also, in the 1990s the growth in direct marketing has led to the use of scorecards to improve the response rate to advertising campaigns. In fact, this was one of the earliest uses in the 1950s when Sears used scoring to decide to whom to send its catalogues (*Lewis 1992*).

Advances in computing allowed other techniques to be tried to build scorecards. In the 80s logistic regression and linear programming, the two main stalwarts of today's card builders, were introduced. More recently, Artificial Intelligence techniques like expert systems and neural networks have been piloted.

At present the emphasis is on changing the objectives from trying to minimize the chance a customer will default on one particular product to looking at how the firm can maximize the profit it can make from that customer. Moreover the original idea of estimating the risk of defaulting has been augmented by scorecards which estimate response (how likely is a consumer to respond to a direct mailing of a new product), usage (how likely is a consumer to use a product), retention (how likely is a consumer to keep using the product after the introductory offer period is over), attrition (will the consumer change to another lender) and debt management (if the consumer starts to become delinquent on the loan how likely are various approaches to prevent default).

2.6 Methods used for Credit Scoring

Originally credit was based on a purely judgmental approach. Credit analysts read the application form and said yes or no. Their decisions tended to be based on the view that what mattered was the 3Cs or the 4Cs or the 5Cs.

Credit scoring nowadays is based on statistical or operational research methods. The statistical tools include discriminant analysis which is essentially linear regression, a variant of this called logistic regression and classification trees, sometimes called recursive partitioning algorithms. The Operational Research techniques include variants of linear programming. Most scorecard builders use one of these techniques or a combination of the techniques. Credit scoring also lends itself to a number of different non-parametric statistical and modelling approaches. Ones that have been piloted in the last few years include the ubiquitous neural networks, expert systems, genetic algorithms and nearest neighbor methods. It is interesting that so many different approaches can be used on the same classification problem.

Part of the reason is that credit scoring has always been based on a pragmatic approach to the credit granting problem. The object is to predict who will default not to give explanations for why they default or answer hypothesis on the relationship between default and other economic or social variables. That is what Capon (*Noel Capon 1982*) considered to be one of the main objections to credit scoring in his critique of the subject.

A sample of previous applicants is taken, which can vary from a few thousand to as high as hundreds of thousands, (not a problem in an industry where firms often have portfolios of tens of millions of customers). For each applicant in the sample, one needs their application form details and their credit history over a fixed period - say 12 or 18 or 24 months.

One then decides whether that history is acceptable, i.e., are they bad customers or not, where a definition of a bad customer is commonly taken to be someone who has missed three consecutive months of payments. There will be a number of customers where it is not possible to determine whether they are good or bad because they have not been customers long enough or their history is not clear. It is usual to remove this set of “intermediates” from the sample.

One question is what is a suitable time horizon for the credit scoring forecast - the time between the application and the good/bad classification. The norm seems to be twelve to eighteen months. Analysis shows that the default rate as a function of the time the customer has been with the organization builds up initially and it is only after twelve months or so (longer usually for loans) that it starts to stabilize.

Thus, any shorter a horizon is underestimating the bad rate and not reflecting in full the types of characteristics that predict default. A time horizon of more than two years leaves the system open to population drift in that the distribution of the characteristics of a population change over time, and so the population sampled may be significantly different from that the scoring system will be used on.

One is trying to use what are essentially cross-sectional models, i.e., the ones that connect two snapshots of an individual at different times, to produce models that are stable when examined longitudinally over time. The time horizon - the time between these two snapshots - needs to be chosen so that the results are stable over time. Another open question is what proportion of goods and bads to have in the sample. Should it reflect the proportions in the population or should it have equal numbers of goods and bads. Henley discusses some of these points in his thesis (*Henley 1995*).

Credit scoring then becomes a classification problem where the input characteristics are the answers to the application form questions and the results of a check with a credit reference bureau and the output is the division into ‘goods’ and ‘bads’. One wants to divide the set of answers A into two subsets $X \epsilon A_N$ the answers given by those who turned out bad, and $X \epsilon A_G$, the set of answers of those who turned out to be good. The rule for new applicants would then be - accepted if their answers are in the set A_G; reject if their answers are in the set A_N. It is also necessary to have some consistency and continuity in these sets and so we accept that we will not be able to classify everyone in the sample correctly. Perfect classification would be impossible anyway since, sometimes, the same set of answers is given by a ‘good’ and a ‘bad’. However, we want a rule that misclassifies as few as possible and yet still satisfy some reasonable continuity requirement.

2.7 Default Prediction Studies

The literature about default prediction methodologies is substantial. Many authors during the last 40 years have examined several possible realistic alternatives to predict customers’ default or business failure. The seminal works in this field were *Beaver (1967)* and *Altman (1968)*, who developed univariate and multivariate models to predict business failures using a set of financial ratios. *Beaver (1967)* used a dichotomous classification test to determine the error rates a potential creditor would experience if he classified firms on the basis of individual financial ratios as failed or non-failed. He used a matched sample consisting of 158 firms (79 failed and 79 non-failed) and he analyzed 14 financial ratios. *Altman (1968)* used a multiple discriminant analysis technique (MDA) to solve the inconsistency problem linked to the Beaver’s univariate analysis and to assess a more complete financial profile of firms. His analysis drew on a matched sample containing 66 manufacturing firms (33 failed and 33 non-failed) that filed a bankruptcy petition during the period 1946-1965. Altman examined 22 potentially helpful financial ratios and ended up selecting five as providing in combination the best overall prediction of corporate bankruptcy. The variables were classified into five standard ratios categories, including liquidity, profitability, leverage, solvency and activity ratios.

MDA is based on two restrictive assumptions:

1. The independent variables included in the model are multivariate normally distributed;
2. The group dispersion matrices (or variance- covariance matrices) are equal across the failing and the non-failing group. See *Karels and Prakash (1987)* and *McLeay and Omar (2000)* for further discussions about this topic. *Zmijewski (1984)* was the pioneer in applying probability analysis to predict default, but, until now, logit analysis has given better results in this field.

For many years thereafter, MDA was the prevalent statistical technique applied to the default prediction models. It was used by many authors (*Deakin (1972), Robert Edmister (1972), Marc Blum (1974), Robert Eisenbeis (1977), Altman et al. (1977), Bilderbeek (1979), Bernard Micha (1984), Gombola et al. (1987), Lussier (1995)).*

However, in most of these studies, authors pointed out that two basic assumptions of MDA are often violated when applied to the default prediction problems 10. Moreover, in MDA models, the standardized coefficients cannot be interpreted like the slopes of a regression equation and hence do not indicate the relative importance of the different variables. Considering these MDA's problems, *Ohlson (1980),* for the first time, applied the conditional logit model to the default prediction's study ll. The practical benefits of the logit methodology are that it does not require the restrictive assumptions of MDA and allows working with disproportional samples. Ohlson used a data set with 105 bankrupt firms and 2,058 non-bankrupt firms gathered from the COMPUSTAT database over the period 1970-1976. He based the analysis on nine predictors (7 financial ratios and 2 binary variables), mainly because they appeared to be the ones most frequently mentioned in the literature. The model's performance, in terms of classification accuracy, was lower than that reported in the previous studies based on MDA (*Altman, 1968 and Altman et al, 1977)*. But reasons were provided as to why logistic analysis was preferable.

From a statistical point of view, logit regression seems to fit well the characteristics of the default prediction problem, where the dependent variable is binary (default/non-default) and with the groups being discrete, non-overlapping and identifiable. The logit model yields a score between zero and one which critics of the logit technique, have pointed out the specific functional form of a logit regression can lead to bimodal (very low or very high) classification and probabilities of default.

Conveniently gives the probability of default of the client 12. Lastly, the estimated coefficients can be interpreted separately as the importance or significance of each of the independent variables in the explanation of the estimated PD. After the work of Ohlson (1980), most of the academic literature (*Zavgren (1983), Becchetti and Sierra (2002))* used logit models to predict default. Despite the theoretic differences between MDA and logit analysis, studies show that empirical results are quite similar in terms of classification accuracy. Indeed, after careful consideration of the nature of the problems and of the purpose of this study, we have decided to choose the logistic regression as an appropriate statistical technique. For comparison purposes, however, we also analyze results using MDA.

Determining the probability of default, PD, in consumer credits, loans and credit cards is one of the main problems to be addressed by banks, savings banks, savings cooperatives and other credit companies. This is a first step needed to compute the capital in risk of insolvency, when their clients do not pay their credits, which is called default. The risk coming from this type of situation is called credit risk, which has been the object of research since the middle of last century. The importance of credit risk, as part of financial risk analysis, comes from the New Basel Capital Accord (Basel II), published in 1999 and revised in 2004 by the Basel Committee for Banking Supervision (BCBS). This accord consists of three parts, called pillars. They constitute a universal theoretical framework for the procedures to be followed by credit companies in order to guarantee minimal capital requirements, called statistical provisions for insolvency (SPI).

Pillar I of the new accord establishes the parameters that play some role in the credit risk of a financial company. These are the probability of default, PD, the exposition after default, EAD, and the loss given default, LGD. The quantitative methods that financial entities can use are those used for computing credit risk parameters and, more specifically, for computing PD. These are the standard method and the internal ratings-based method (IRB). Thus, credit companies can elaborate and use their own credit qualification models and, by means of them, conclude the Basel implementation process, with their own estimations of SPI.

There is an extensive literature on quantitative methods for credit risk, since the classical Z-score model introduced by *Altman (1968)*. Nowadays there exist plenty of approaches and perspectives for modelling credit risk starting from PD. Most of them have provided better predictive powers and classification error rates than Altman's discriminant model, for credit solicitors (application scoring), as well as for those who are already clients of the bank (behavioral scoring). This is the case of logistic regression models, artificial neural networks (ANN), support vector machines (SVM), as well as hybrid models, as mixtures of parametric models and SVM.

The idea of using survival analysis techniques for constructing credit risk models is not new. It started with the paper by *Narain (1992)* and, later, was developed over the years. A common feature of all these papers is that they use parametric or semiparametric regression techniques for modelling the time to default (duration models), including exponential models, Weibull models and Cox's proportional hazards models, which are very common in this literature. The model established for the time to default is then used for modelling PD or constructing the scoring discriminant function.

2.8 Linear Regression

Ordinary linear regression (Reg) is the simplest, compared all other techniques. Using the dummy variable for the dependent variable of good/bad indicator (say, define it as 1 if borrower is good, 0 if he/she is bad) and regressing it on a set of characteristics of borrowers by the standard least square approach will produce the estimated "probability" of being good. It is well known as the Linear Probability Model. Its main drawback is that there is no guarantee that the estimated probability would happen within the interval of [1, 0]. *Orgler (1970)* has applied regression analysis in a model for commercial loans and *Orgler (1971)* used it for evaluating existing loans.

2.9 Discriminant Analysis

Discriminant analysis (DA) is a technique for first identifying the "best" characteristics of the debtors, known as discriminator variables, which provide the maximum discrimination between high and low default risk borrowers. Generally, assumption of multivariate normality of the variables is required. *Durand (1941)* considered the use of discriminant analysis for the scoring system. Another account of its application in credit scoring is given by *Myers and Forgy (1963).*

2.10 Decision Tree Rule

The Decision Tree (D. tree), also known as the classification tree and recursive partitioning, tries to split the population into two sub-groups which are more homogeneous by making use of the possible characteristics of the debtors. It keeps applying this procedure until one has a number of groups identified as either good or bad debtors. Application of such a method in credit scoring is given in *Makowski (1985)* and *Mehta (1968).*

2.11 K-Nearest Neighbor Classifiers

K-Nearest Neighbor (KNN) classifiers classify a data instance (i.e., borrower) by considering only the k-most similar data instances. The class label is then assigned according to the class of the majority of the k-nearest neighbors. To measure the distances among the data instances, it is common to choose the Euclidean distance, in which the characteristics of the borrower are taken into account. This approach was applied in the credit-scoring context by *Chatterjee and Barcun (1970)* and *Henley and Hand (1996).*

2.12 Support Vector Machines

Support Vector Machine (SVM) is closely related to linear programming. The major difference is that it not only minimizes the value of errors, but also maximizes the marginal difference between the good ones and the bad ones. As a Bayesian network classifier, SVM is uncommon in credit scoring literature. Recently, *Baesens et al. (2003)* found that SVM performs very well.

2.13 Neural Networks

Neural networks (NN) are mathematical representations inspired by the functioning of the human brain. As *Hand and Henley (1997)* stated, the type of NN that is normally applied to credit scoring problems can be viewed as a statistical model involving linear combinations of nested sequences of non-linear transformations of linear combinations of variables. In brief, NN can be considered as a form of non-linear regression. *Rosenberg and Glei (1994)* described applications of NN to corporate credit decisions and fraud detection and *Davies et al. (1992)* compared it with other scorecards.

2.14 Probability Density Function of Credit Losses

When estimating the amount of economic capital needed to support their credit risk activities, many large sophisticated banks employ an analytical framework that relates the overall required economic capital for credit risk to their portfolio's ***probability density function*** of credit tosses

(PDF), which is the primary output of a credit risk model. Exhibit 1 illustrates this relationship. A bank would use its credit risk modelling system to estimate such a PDF.

An important property of a PDF is that the probability of credit losses exceeding a given amount X (along the x-axis) is equal to the (shaded) area under the PDF to the right of X. A risky portfolio is one whose PDF has a relatively long and fat tail. The expected credit loss (shown as the left-most vertical line) shows the amount of credit loss the bank would expect to experience on its credit portfolio over the chosen time horizon. Banks typically express the risk of the portfolio with a measure of *unexpected credit loss* (i.e., the amount by which actual losses exceed the expected loss) such as the standard deviation of losses or the difference between the expected loss and some selected target credit loss quantile.

The estimated economic capital needed to support a bank's credit risk exposure is generally referred to as its required economic capital for credit risk. The process for determining this amount is analogous to value at risk (VaR) methods used in allocating economic capital against market risks.

Specifically, the economic capital for credit risk is determined so that the estimated probability of unexpected credit loss exhausting economic capital is less than some target insolvency rate.

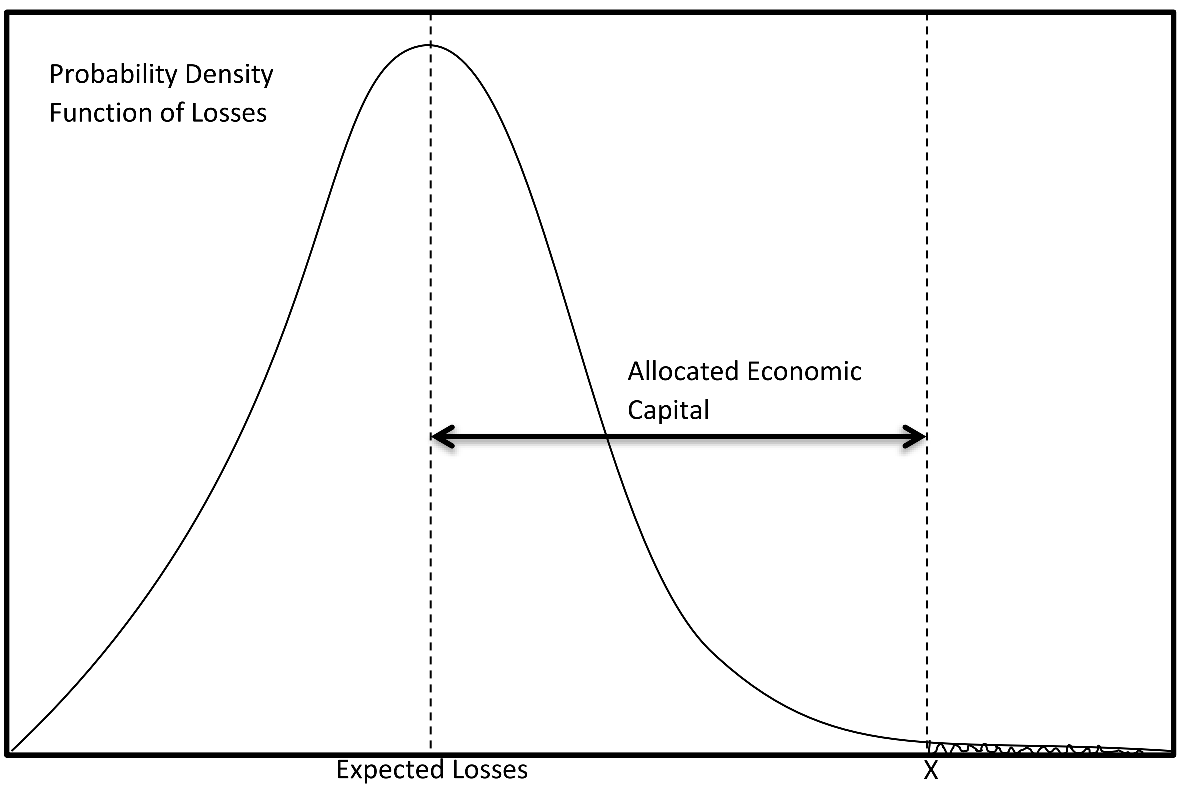


Exhibit 1

Capital allocation systems generally assume that it is the role of reserving policies to cover expected credit losses, while it is that of economic capital to cover unexpected credit losses

Thus, required economic capital is the additional amount of capital necessary to achieve the target insolvency rate, over and above needed for coverage of expected losses.

In Exhibit 1:

1. For a target insolvency rate equal to the shaded area, the required economic capital equals the distance between the two dotted lines.
2. Broadly defined, a credit risk model encompasses all of the policies, procedures and practices used by a bank in estimating a credit portfolio's PDF.

2.15 Survival Modelling

Because of the limitation of Classification Models, the credit scoring model is extended to estimate the time-to-default, instead of whether the borrower will default or not. As suggested by *Banasik et al. (1999)*, it has now become important to know not only if but also when the borrower would default. It is similar to the ideas of survival analysis in mortality and equipment reliability. Since *Narain (1992)* applied survival analysis for credit scoring, it has been widely investigated for credit risk management. *Banasik et al. (1999)* analyze the time-to-default and time-to-early-repayment by semi-parametric proportional hazards model (Cox model) and two parametric proportional hazards models (with exponential and Weibull baseline hazards).

Stepanova and Thomas (2002) adopt the Cox model to personal loan data by coarse-classifying of characteristics and by including interactions of time-by-characteristics. *Stepanova and Thomas (2001)* further develop survival analysis techniques in credit risk modelling by estimating the expected profit of personal loans. Most of them not only estimate the probability of default of the loan over time, but also classify the borrowers into either “good” or “bad”. Concerning the accuracy of classification, the survival analysis is comparable with logistic regression, the most common approach for credit risk modelling (*Thomas et al. 2002*)

2.16 General Issues in Credit Risk Modelling

The field of credit risk modelling has developed rapidly over the past few years to become a key component in the risk management systems at financial institutions. In fact, several financial institutions and consulting firms are actively marketing their credit risk models to other institutions. In essence, such models permit the user to measure the credit risk present in their asset portfolios. This information can be directly incorporated into many components of the user's credit portfolio management, such as pricing loans, setting concentration limits and measuring risk-adjusted profitability (*Stepanova and Thomas 2002*).

As summarized by the Federal Reserve System Task Force of USA on Internal Credit Risk Models (FRSTF, 1998) and the Basle Committee on Banking Supervision (BCBS, 1999), there exists a wide variety of credit risk models that differ in their fundamental assumptions, such as their definition of credit losses; i.e., default models define credit losses as loan defaults, while market-to-market or multi-state models define credit losses as rating migrations of any magnitude.

However, the common purpose of these models is to forecast the probability distribution function of losses that may arise from a bank's credit portfolio (*Stepanova and Thomas 2002*). Such loss distributions are generally not symmetric. Since credit defaults or rating changes are not common events and since debt instruments have set payments that capture possible returns, the loss distribution is generally skewed toward zero with a long right-hand tail.

Although an institution may not use the entire loss distribution for decision-making purposes, credit risk models typically characterize the full distribution. A credit risk model's loss distribution is based on two components: the multivariate distribution of the credit losses on all the credits in its portfolio and a weighting vector that characterizes its holdings of these credits. This ability to measure credit risk clearly has the potential to greatly improve banks' risk management capabilities. With the forecasted credit loss distribution in hand, the user can decide how best to manage the credit risk in a portfolio, such as by setting aside the appropriate loan loss reserves or by selling loans to reduce risk. Such developments in credit risk management have led to suggestions, such as by International Swaps and Derivatives Association, ISDA (1998) and Institute of International Finance, IIF (1998) that bank regulators permit, as an extension to risk-based capital standards, the use of credit risk models for determining the regulatory capital to be held against credit losses. Currently, under the Basle Capital Accord, regulated banks must hold 8% capital against their risk-weighted assets, where the weights are determined according to very broad criteria. For example, all corporate loans receive a 100% weight, such that banks must hold 8% capital against such loans.

Proponents of credit risk models for regulatory capital purposes argue that the models could be used to create risk- weightings more closely aligned with actual credit risks and to capture the effects of portfolio diversification. These models could then be used to set credit risk capital requirements in the same way that VaR models are used to set market risk capital requirements under the MRA.

However, as discussed by FRSTF (1998) and BCBS (1999), two sets of important issues must be addressed before credit risk models can be used in determining risk-based capital requirements. The first set of issues corresponds to the quality of the inputs to these models, such as accurately measuring the amount of exposure to any given credit and maintaining the internal consistency of the chosen credit rating standard.

CHAPTER THREE

METHODOLOGY

INTRODUCTION

Researchers use the survival analysis techniques in a variety of contexts that share a common characteristic. The interest concentrates on describing whether or when events occur. It is necessary to use the survival analysis techniques

3.1 Research Design

The study is a financial-institution based retrospective case study research design. The focus is on personal loans whose maturity is three years and above. The performance of accounts was observed for a period of 36 months from January 1, 2019 to January 1, 2022. In the context of this study, number censored included those who made early settlement. The study considered loans taken within the month of January 2019.

3.2 Target population

The target population is male and female applicants who took personal loans. In survival analysis, one must consider a key analytical problem called censoring. In essence, censoring occurs when we have some information about an individual's survival time, but do not know the exact survival time. There are a number of types of censoring, such as random, interval, left, and right censoring. In credit scoring application, most of the cases are right censoring. For example, suppose we follow a group of borrowers for 3 years. If we observe borrower A fails to repay at 15th month, he is certainly classified as a default case and his default time is 15. On the other hand, consider borrower B, who repays on time during the whole observed period. We do not know his exact default time but are sure that it must be greater than 36. For such case, borrower B is known as a right censored observation. Another example of right censoring could be when borrower C repays on time from the 1st month to the 12th month. At the 12th month, we do not have future repayment pattern of borrower C. As borrower B, we do not know the exact default time of borrower C, we only know that it must be greater than 12. This is also a right censoring example.

3.3 Sample size

The data set consisted of application information 100 male and 100 female applicants, a total of 200 successful personal loan applicants randomly picked from the banks database together with the repayment status of each month of the observed period.

3.4 Data collection

Method of data collection is through observation and documents. Secondary data will be collected from one of the leading commercial banks in Kenya.

3.5 Descriptive methods of Time-to-event

Survival analysis is a statistical method for modelling the time to some events for a population of individuals. For example, events may refer to death in medical application, or recidivism of released prisoners in criminology application, or first bought of a new product by customer in marketing studies. The time to the occurrence is termed as survival time or lifetime. In application to credit risk modelling, the events refer to default of a loan and therefore its lifetime refers to time-to-default T.

Default times are subject to random variation and are thus random variables. To describe their randomness, there are five standard ways:

Probability Density function (PDF), $f(t) = \lim_{\Delta t \to 0} \frac{P(t \le T \le t+\Delta t)}{\Delta t}$

Cumulative Distribution Function (CDF), $F(t) = \int_0^t f(u)du = P(T \le t)$

Survival Function, $S(t) = \int_t^{\infty} f(u)du = 1 - F(t) = P(T \ge t)$

Hazard Function, $\lambda(t) = \lim_{\Delta t \to 0} \frac{P(t \le T \le t+\Delta t / T \ge t)}{\Delta t}$

$$\frac{f(t)}{S(t)} = \frac{-d \ln S(t)}{dt}$$

Cumulative hazard function, $H(t) = \int_0^{\infty} \lambda(t)dt = -\ln S(t)$

These five formulations are mathematically equivalent but they highlight different aspects of the default time. The distribution function tells us the probability that default occurs at or before time *t*. Conversely, survivor function is the probability that default does not occur at or before time *t*; in other words, the loan survives (non-default), at least, to time *t*.

The interpretation of hazard function is slightly tricky. It is the “rate” that borrower defaults at time t, conditional on his staying on the books up to that time. Note that hazard is not a probability and thus can be greater than one.

3.6 Model Derivation

Estimating the Survival function.

Recall that,

If we assume the exponential distribution, which assumes that an applicant's likelihood of experiencing the event (defaulting) is independent of how long that person has been event-free.

Then the survival function is given by,

$$S(t) = P(T \geq t) = e^{-\lambda t}$$

We estimate the hazard within the interval containing event time j as:

$$\lambda_j = \frac{D_j}{N_j}$$

And $\lambda_j = 0$ if $D_j = 0$.

Kaplan-Meier estimator can be derived from maximum likelihood estimation of hazard function.

Given D_j as the number of events (defaults) and N_j as the total applicants at risk of defaulting during interval $j; (1 \leq j \leq k)$, discrete hazard rate λ_j can be defined as the probability of an applicant with an event during interval j.

The survival rate can be written

$$S(t) = \prod_{j:j \leq k} (1 - \lambda_j)$$

The likelihood function for the hazard function up to time interval j is;

$$\mathcal{L}\left(\lambda_{j:j \leq k} \middle| D_{j:j \leq k}, N_{j:j \leq k}\right) = \prod_{j=1}^{k} \lambda_j^{D_j} (1 - \lambda_j)^{N_j - D_j}$$

Therefore, the log-likelihood will be;

$$\log(\mathcal{L}) = \sum_{j=1}^{k} (D_j \log \lambda_j + (N_j - D_j) \log(1 - \lambda_j))$$

Finding the maximum of log likelihood with respect to λ_j yields:

$$\frac{\partial \log(\mathcal{L})}{\partial \lambda_j} = \frac{D_j}{\widehat{\lambda_j}} - \frac{N_j - D_j}{1 - \hat{\lambda}_j} = 0$$

$$\Longrightarrow \widehat{\lambda_j} = \frac{D_j}{N_j}$$

Where $\widehat{\lambda_j}$ is used to denote maximum likelihood estimate. Given this result, we can state;

$$\hat{S}(t) = \prod_{j:j\leq k} (1 - \widehat{\lambda_j}) = \prod_{j:j\leq k} (1 - \frac{D_j}{N_j})$$

3.7 Data entry and analysis

Suppose we observe a population of N applicants in the presence of non-informative right censoring, and we observe D defaults;

$t -$ time.

j – time interval, $(t_{j-1}\ to\ t_j)$

N_j – number of applicants who did not default and considered at risk during interval j.

D_j – number of defaults at interval j.

C_j – number of applicants who are censored during interval j.

λ_j – hazard function/proportion of defaulting during interval j (or suffering event) $\{D_j/N_j\}$

p_j – proportion of surviving (not defaulting) during interval j (remaining event free) $\{1 - \lambda_j\}$

S_t – the cumulative survival probability during interval $j(t_{j-1}\ to\ t_j)$\{The probability of applicants surviving past t=0 is $S_0 = 1$, all applicants are event free at start of the study.

Proportion of surviving past each subsequent interval is computed using principles of conditional probability. i.e.

$S_1 = p_1$, probability applicant survives past interval 1.

$S_2 = p_2 \times S_1$, probability applicant survives past interval 2.

In general, $S_{t+1} = p_{t+1} \times S_t\}$

F_t – the cumulative failure/default probability during interval t $\{1 - S_t\}$

Let $t_0 < t_1 < t_2 \ldots \ldots .. < t_k$ be ordered times at which defaults were observed and k is the total lifespan of the study (k=36). We will take t_0 to t_1 to be interval 1, $(j = 1)$,t_1 to t_2 to be interval 2, $(j = 2)$, and so on and so forth.

More defaults can occur at a single interval therefore, we do not assume that $k = D$.

Suppose that D_j defaults were observed at interval j $(1 \leq j \leq k)$;

So that $D_1 + D_2 + \cdots \ldots \ldots + D_k = D$.

The censored observations are $N - D$.

Suppose that C_j lives are censored at interval $j(1 \leq j \leq k)$, where we define $t_0 = 0$ and $t_{(k+1)} = \infty$ to allow for censored observations after the last observable time; then:

$C_1 + C_2 + \cdots \ldots \ldots \ldots + C_k = N - D$.

C_j = applicants removed from the investigation at interval $j(1 \leq j \leq k)$.

The hazard of defaulting at any particular interval, j, when an event takes place is equal to D_j/N_j where D_j is the number of applicants who defaulted at interval j and N_j is the average risk set at that interval.

Log rank test is used to test the null hypothesis of no difference in survival between the two groups. We will use one that is closely linked to the chi-square test statistic and compares observed to expected numbers of events at each time point over the follow-up period.

We use the following;

$$x^2 = \sum \frac{\left(\sum O_{jt} - \sum E_{jt}\right)^2}{\sum E_{jt}}$$

To compare two or more groups, we need the risk set (N), observed (O) and the expected (E) values of each group.

And

$\sum O_{jt} -$ sum of the observed defaults in the j^{th} group over time.

$\sum E_{jt} -$ sum of expected defaults in the j^{th} group over time.

Where $j = 1,2$; $1 = male\ applicants, 2 = female\ applicants$.

And;

H_0 : the two survival curves are identical ($S_{1t} = S_{2t}$) vs

H_1 : the two survival curves are not identical ($S_{1t} \neq S_{2t}$)

The log rank statistic has degrees of freedom equal to $k - 1$, where k represents the number of groups. In this case, $k = 2$, so the test has 1 degree of freedom.

3.8 Reasons why we use Kaplan-Meier model

1. It makes interpreting the output easier.
2. It is commonly used to describe survival rate.
3. It is used to compare two study population.
4. It is an intuitive graphical presentation.

CHAPTER FOUR

RESEARCH FINDINGS AND INTERPRETATION OF RESULTS

INTRODUCTION

In this chapter, we are going to discuss research findings and explanation of results in the context of the research area.

4.1 Research Findings

After keen consideration and selection of applicants from the banks database, the following results were obtained for analysis.

Number of applicants who made early settlements;

Males: 48 applicants

Females: 43 applicants

Number of applicants who defaulted;

Male: 52 applicants

Female: 57 applicants

Time in months at which borrowers made early settlement on their loans or defaulted are as follows;

Early settlements:

Male:4,4,6,7,7,8,9,10,10,11,12,13,13,15,16,16,17,17,17,17,18,18,18,19,19,20,21,21,22,22,22,23,23,24,24,25,26,27,28,28,30,30,31,32,33,33,35,35

Female:4,8,10,10,10,11,11,11,12,12,13,13,13,14,15,17,18,19,19,19,20,21,22,22,22,22,23,24,24,24,25,26,27,28,28,29,29,30,31,33,33,33,34

Defaults:

Male;3,3,4,5,5,5,6,6,6,7,7,7,8,8,9,9,9,9,9,10,10,11,11,11,11,12,12,13,14,15,15,15,16,16,17,19,19,20,21,22,22,22,24,24,26,27,27,28,29,29,30,31

Female:2,4,5,5,6,7,7,7,8,8,8,10,10,10,11,11,12,12,12,12,12,14,15,15,15,15,16,16,17,17,17,17,17,18,19,19,20,21,23,23,23,24,26,27,28,28,28,29,29,29,30,31,31,33,33,34,35

4.2 Model Application

This data was processed using R Studio and Excel, and the model output are as follows:

4.3 Survival Analysis for Male Applicants

For gender factor 1 (Male):

Table 4.1:

Male applicant's data						
Time In months t	Number at risk N_j	Number of defaults D_j	Number censored C_j	Hazard rate $\boldsymbol{\lambda}_j$	Cumulative Survival Probability S_t	Cumulative Failure Probability $1 - S_t$
0	100	0	0		1	0
3	100	2	0	0.020	0.980	0.020
4	97	1	2	0.010	0.970	0.030
5	94	3	0	0.032	0.939	0.061
6	90	3	1	0.033	0.908	0.092
7	85	3	2	0.035	0.876	0.124
8	82	2	1	0.024	0.854	0.146
9	76	5	1	0.066	0.798	0.202
10	72	2	2	0.028	0.776	0.224
11	67	4	1	0.060	0.730	0.270
12	64	2	1	0.031	0.707	0.293
13	61	1	2	0.016	0.695	0.305
14	60	1	0	0.017	0.684	0.316
15	56	3	1	0.054	0.647	0.353
16	52	2	2	0.038	0.622	0.378
17	47	1	4	0.021	0.609	0.391
18	44	0	3		0.609	0.391
19	40	2	2	0.050	0.578	0.422
20	38	1	1	0.026	0.563	0.437
21	35	1	2	0.029	0.547	0.453
22	29	3	3	0.103	0.491	0.509
23	27	0	2		0.491	0.509
24	23	2	2	0.087	0.448	0.552
25	22	0	1		0.448	0.552
26	20	1	1	0.050	0.425	0.575
27	17	2	1	0.118	0.375	0.625
28	14	1	2	0.071	0.349	0.651
29	12	2	0	0.167	0.290	0.710
30	9	1	2	0.111	0.258	0.742
31	7	1	1	0.143	0.221	0.779
32	6	0	1		0.221	0.779
33	4	0	2		0.221	0.779
35	2	0	2		0.221	0.779

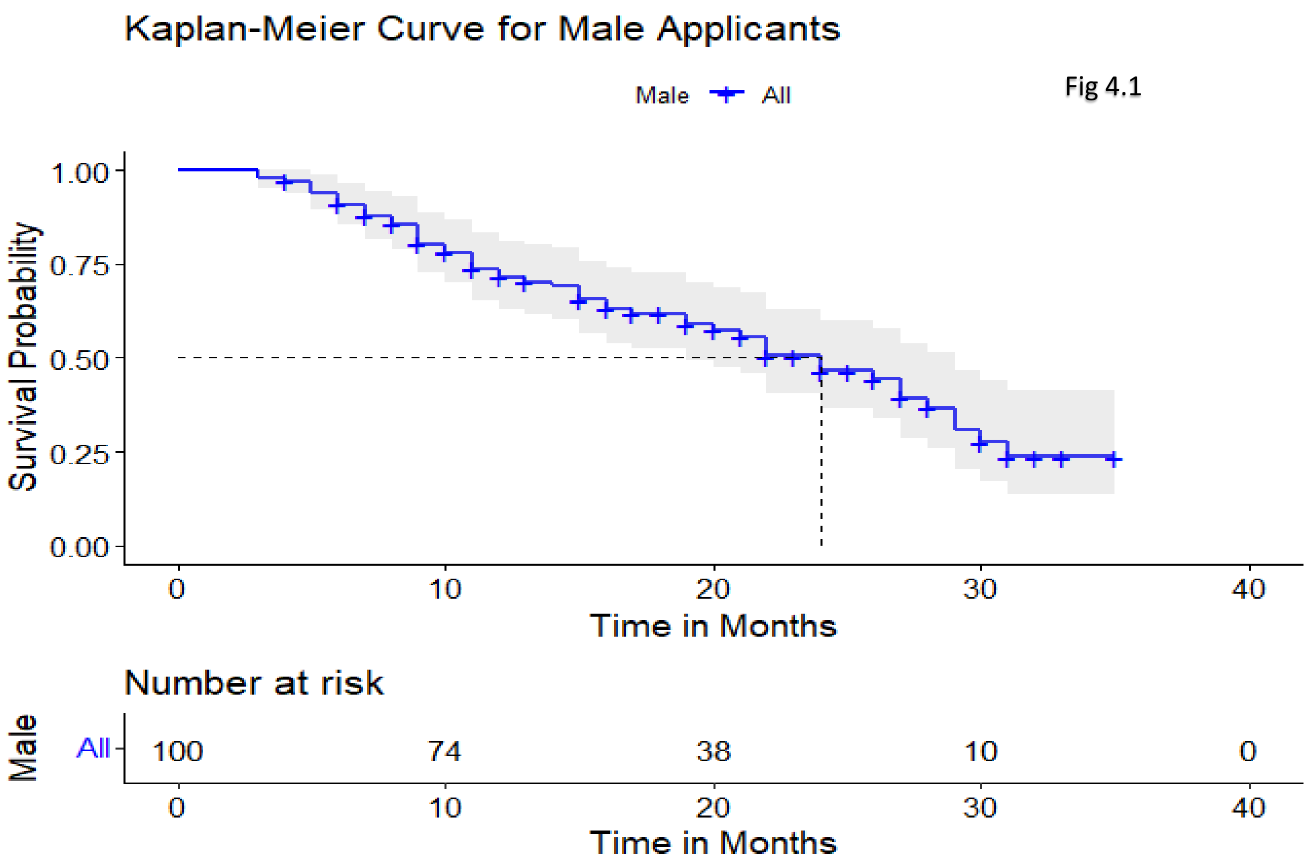

Kaplan-Meier Curve for Male Applicants
Male All
Survival Probability
1.00
0.75
0.50
0.25
0.00
0
10
20
30
40
Time in Months
Number at risk
Male
All
100
74
38
10
0
0
10
20
30
40
Time in Months

Fig 4.1

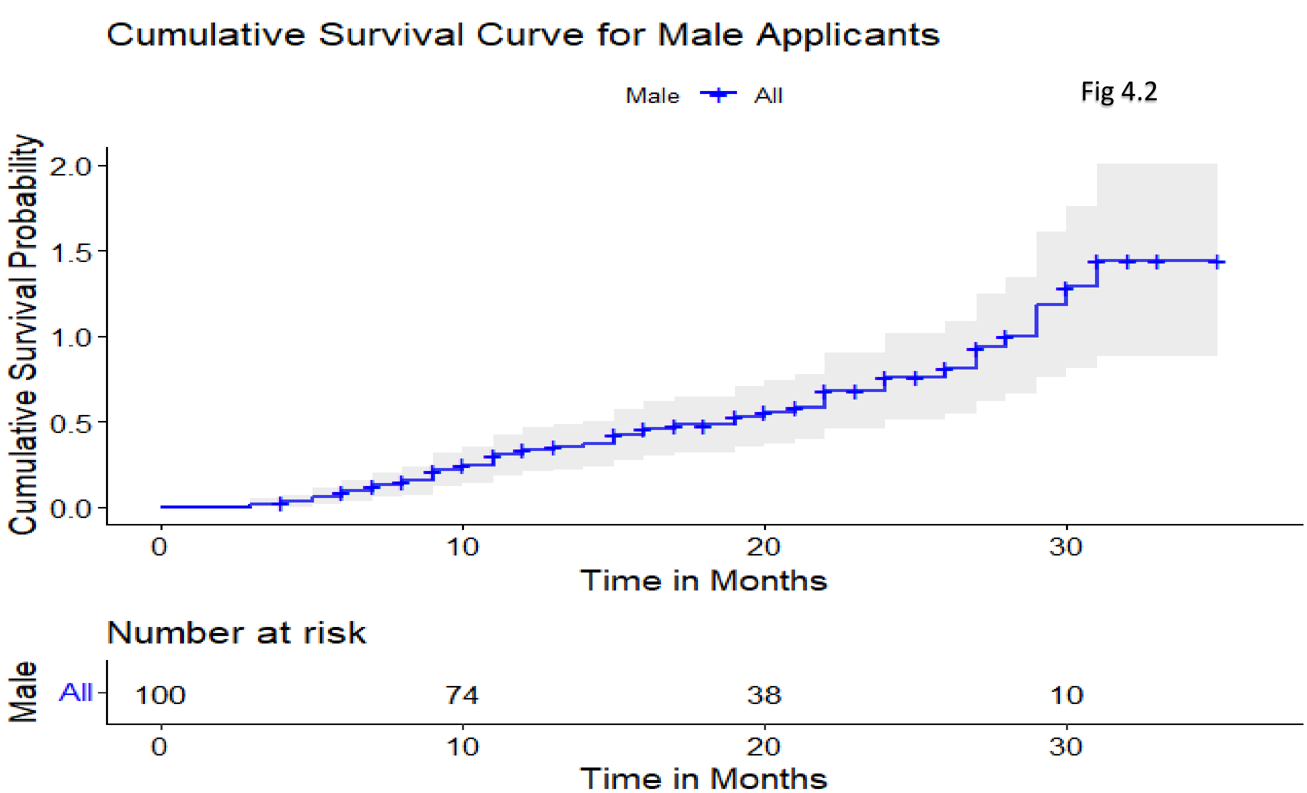

Cumulative Survival Curve for Male Applicants
Male All
Cumulative Survival Probability
2.0
1.5
1.0
0.5
0.0
0
10
20
30
Time in Months
Number at risk
Male
All
100
74
38
10
0
10
20
30
Time in Months

Fig 4.2

```
Number of defaults = 52     censored = 48        (48%)

                mean   median   95% confidence interval
Group = 1        15      24      ( 19       ,       29 )
```

4.4 Survival Analysis for Female Applicants

For gender factor 2 (Female):

Table 4.2:

Female applicant's data						
Time in months t	Number at risk N_j	Number of defaults D_j	Number censored C_j	Hazard Rate $\boldsymbol{\lambda}_j$	Cumulative Survival Probability S_t	Cumulative Failure Probability $1 - S_t$
0	100	0	0		1	0
2	100	1	0	0.010	0.990	0.010
4	99	1	1	0.010	0.980	0.020
5	97	2	0	0.021	0.960	0.040
6	95	1	0	0.011	0.950	0.050
7	94	3	0	0.032	0.919	0.081
8	91	3	1	0.033	0.889	0.111
10	87	4	2	0.046	0.848	0.152
11	81	2	3	0.025	0.827	0.173
12	76	5	2	0.066	0.773	0.227
13	69	0	3		0.773	0.227
14	66	1	1	0.015	0.761	0.239
15	64	4	1	0.063	0.714	0.286
16	59	2	0	0.034	0.689	0.311
17	57	5	1	0.088	0.629	0.371
18	51	1	1	0.020	0.617	0.383
19	49	2	3	0.041	0.591	0.409
20	44	1	1	0.023	0.578	0.422
21	42	1	1	0.024	0.564	0.436
22	40	0	4		0.564	0.436
23	36	3	1	0.083	0.517	0.483
24	32	1	3	0.031	0.501	0.499
25	28	0	1		0.501	0.499
26	27	1	1	0.037	0.482	0.518
27	25	1	1	0.040	0.463	0.537
28	23	3	2	0.130	0.403	0.597
29	18	3	2	0.167	0.336	0.664
30	13	1	1	0.077	0.310	0.690
31	11	2	1	0.182	0.253	0.747
32	8	1	0	0.125	0.222	0.778
33	7	1	3	0.143	0.190	0.810
34	3	1	1	0.333	0.127	0.873
35	1	0	1		0.127	0.873

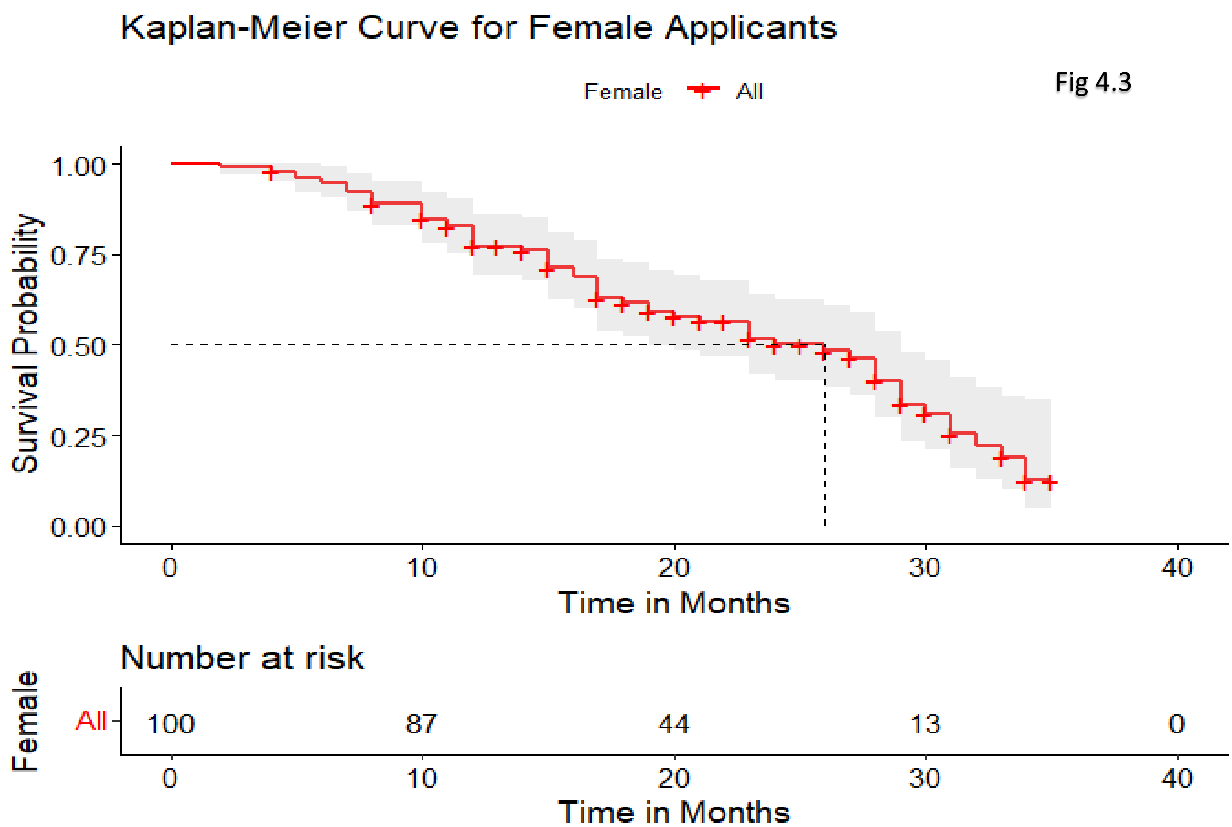

Kaplan-Meier Curve for Female Applicants
Female
All
Survival Probability
1.00
0.75
0.50
0.25
0.00
0
10
20
30
40
Time in Months
Number at risk
Female
All
100
87
44
13
0
0
10
20
30
40
Time in Months

Fig 4.3

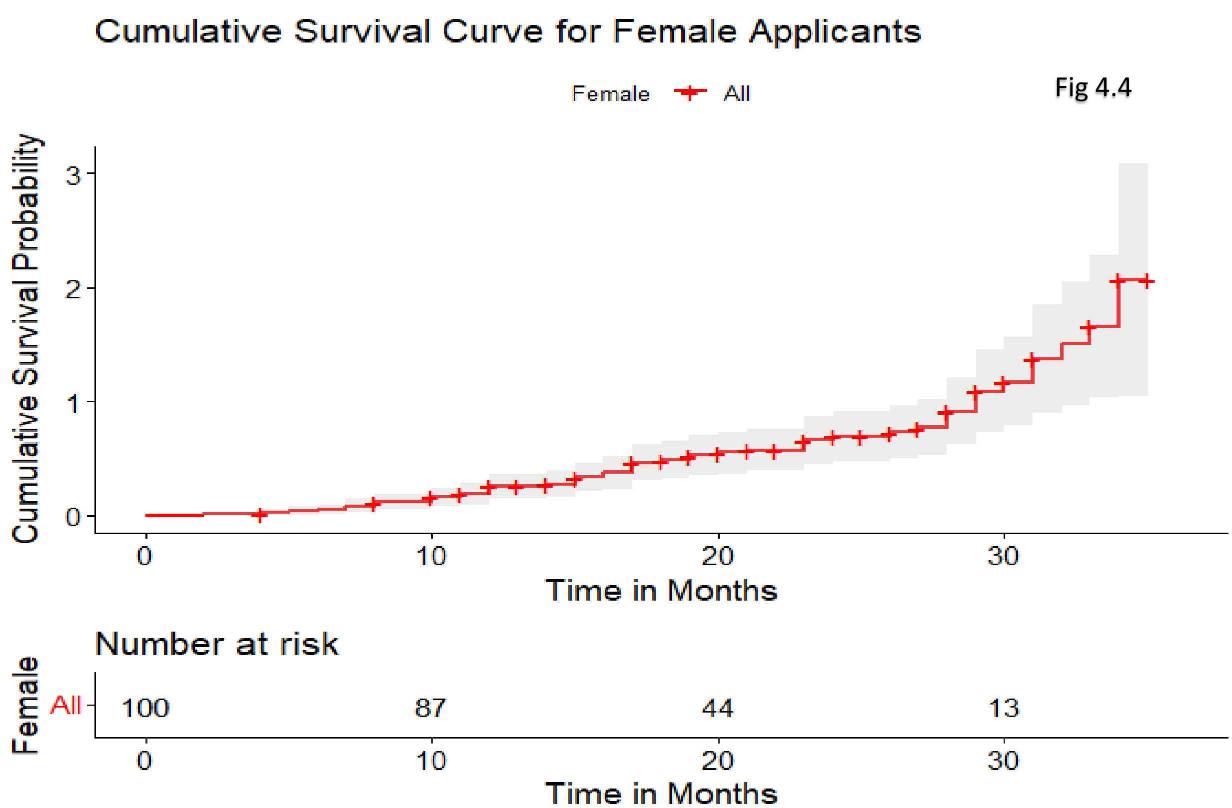

Cumulative Survival Curve for Female Applicants
Female
All
Cumulative Survival Probability
3
2
1
0
0
10
20
30
Time in Months
Number at risk
Female
All
100
87
44
13
0
10
20
30
Time in Months

Fig 4.4

```
Number of defaults = 57     censored = 43        (43%)

                mean     median    95% confidence interval
Group = 2        18        26       ( 19       ,       29 )
```

The following summary was also generated:

```
Combined output:
```

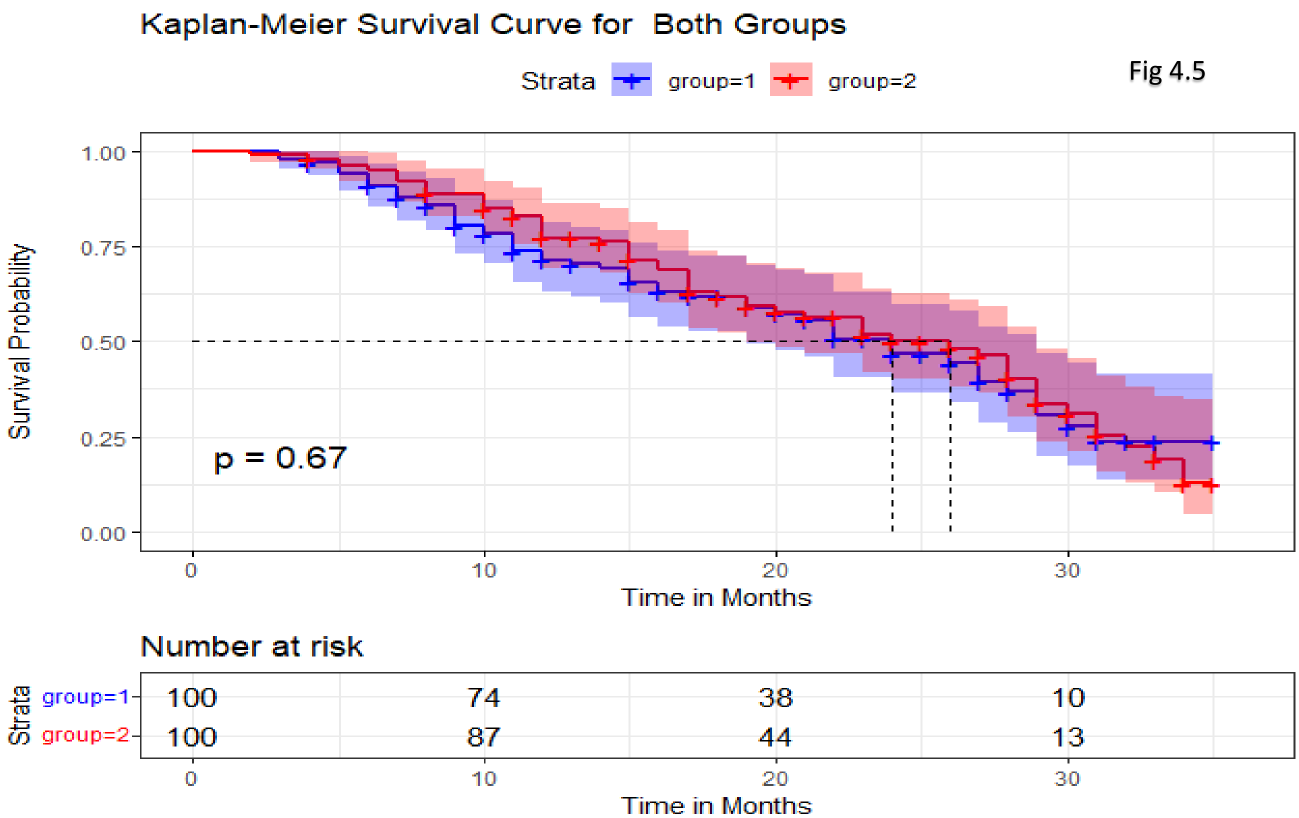
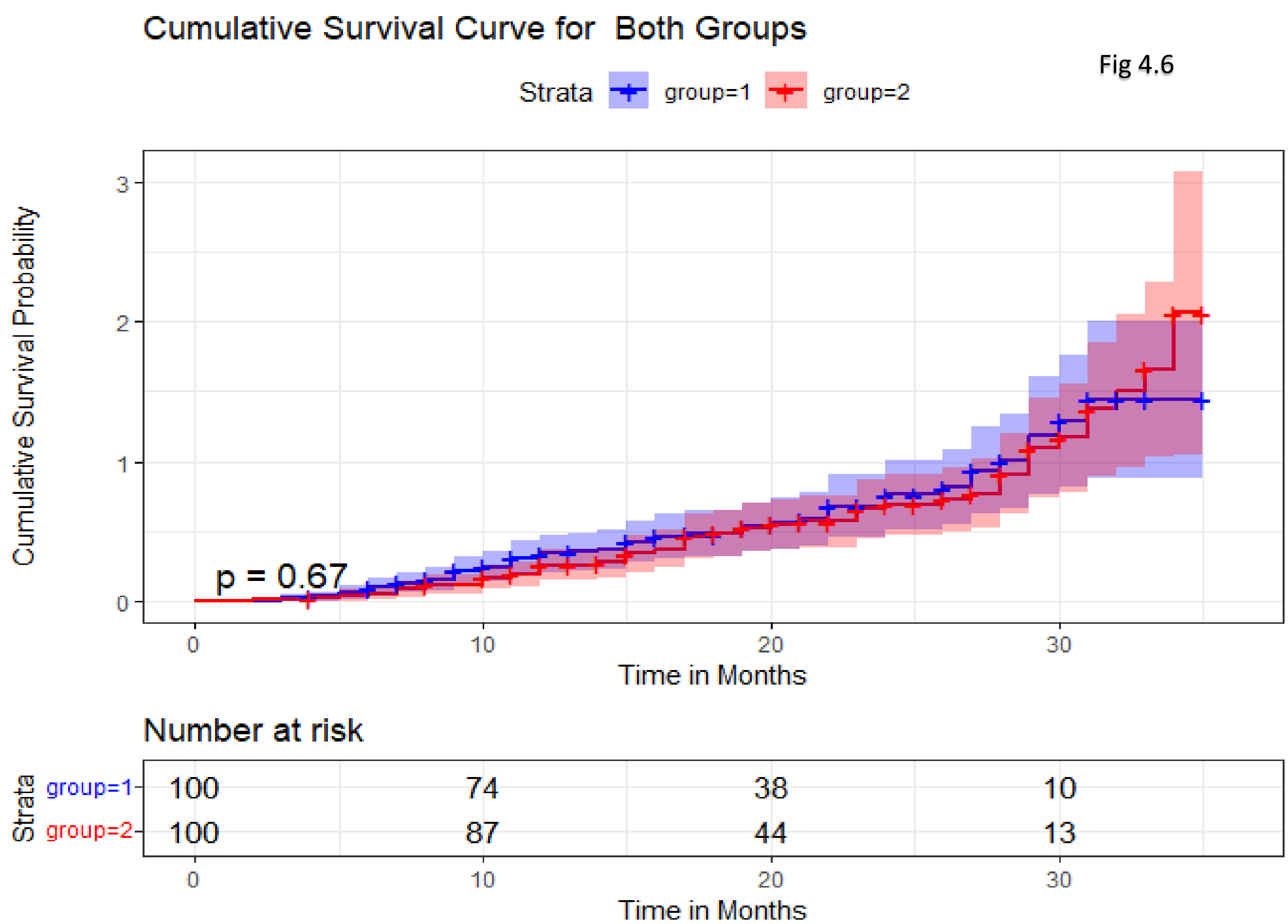


Fig 4.5

Fig 4.6



	Total	Events	Censored	% Censored
Group = 1	100	52	48	48
Group = 2	100	57	43	43
	200	109	91	45.5

Log-rank Test Statistics for Equality of Survival Distribution for Gender

	N	Observed	Expected	(O-E)^2/E	(O-E)^2/V
group=1	100	52	49.8	0.0934	0.18
group=2	100	57	59.2	0.0787	0.18

Chi-square distribution = 0.2, on 1 degrees of freedom, p = 0.7

Group 1 represents male applicants and group 2 represents female applicants. Event 1 denotes default while event 0 represents censored observations.

The survival data on group 1 implies that out of the 100 male applicants for loans maturing in 36 months, 52 defaulted and 48 settled their loan obligations before maturity. Mean survival time of 15 which means a male applicant will take 15 months to default on average. Same interpretation can be attributed to data on female applicants.

By observation of survival curves, it can be seen that the two curves are similar. This is confirmed by the log-rank test statistic ($p = 0.67$) which is greater than 0.05. This shows that the two survival distributions are statistically the same. Thus, it would not be meaningful to classify borrowers on the basis of gender.

CHAPTER FIVE
CONCLUSIONS AND RECCOMENDATIONS

INTRODUCTION

In this chapter, we give the position arrived at based on the research outcome and relates the outcome to the world of practice. Room for further research is also proposed.

5.1 Conclusions

The research findings show that there is no significant difference between the male and female borrowers in terms of their time to default on loan obligations thus the null hypothesis is true. This implies that gender does not affect credit risk. Mean survival times would be a great reference to guide credit granting process on the average maturity for loans that may minimize on default losses and optimize profitability.

5.2 Recommendations and Suggestions for further research

This method of credit risk modelling is quite reliable as it does not make assumptions about loan default distribution unlike parametric methods. However, given that Kaplan-Meier approach is a univariate method, it may be more informative to adopt multivariate techniques like Cox Proportional model to model credit risk. Thus, further research can be conducted on the same data set using other survival techniques.

REFERENCES

Altman, E. I., Haldeman, R. G., & Narayanan, P. (1977). *ZETA Analysis: A New Model to Identify Bankruptcy Risk of Corporations. Journal of Banking and Finance*, 1, 29-54. https://doi.org/10.1016/0378-4266(77)90017-6

Baesens, B., T. Van Gestel, S., Viaene, M., Stepanova, J. & Suykens, J. V. (2003). *Benchmarking state-of-the-art classification algorithms for credit scoring. Journal of the Operational Research Society* 54 627-635. http://dx.doi.org/10.1057/palgrave.jors.260154

Banasik, J., J. N. Crook, L. C. Thomas. 1999. *Not if but when will borrowers default. Journal of the Operational Research Society* 50(12) 1185-1190.

Beaver, W. (1967). *Financial Ratio as Predictors of Failure, Empirical Research in Accounting: Selected Studies 1966. Journal of Accounting Research*, 4, 71-111. https://doi.org/10.2307/2490171

Becchetti, L. and Sierra, J., 2002. Finance, Investment and Innovation: Empirical and Theoretical Challenges (EIFC Working Paper, No. 02-9). *Maastricht: Selbstverlag des Forschungsverbundes European Integration, Financial Systems and Corporate Performance (EIFC)*

Bernard Micha 1984. *Analysis of business failures in France. Journal of Banking & Finance.* Volume 8, Issue 2, Pages 281-291. https://www.sciencedirect.com/journal/journal-of-banking-and-finance

Bilderbeek, J. (1979). *An empirical study of the predictive ability of financial ratios in the Netherlands*. *Zeitschrift Fur Betriebswirtschaft*. 49.5 388-407

Boyle, M., J. N. Crook, R. Hamilton, L. C. Thomas. 1992. *Methods for credit scoring applied to slow payers*. L. C. Thomas, J. N. Crook, D. B. Edelman, ed. *Credit Scoring and Credit Control. Oxford University Press*: Oxford, 75-90.

Chatterjee, S., Barcun, S. (1970). *A nonparametric approach to credit screening. Journal of the American Statistical Association* 65 150-154. http://dx.doi.org/10.2307/2283581

Churchill Gilbert A. Jr., Nevin John R., And Richard Watson R. (1977). *The Role of Credit Scoring in the Loan Decision.* The Credit World. https://doi.org/10.1177/002224298204600209

D.W. Johnson and R.T. Johnson (1992). *Positive Interdependence. Key to Effective Cooperation. Interaction in cooperative groups,* New York: Cambridge University Press.

David Hand & M. G. Kelly 2001. *Lookahead Scorecards for New Fixed Term Credit Products. Journal of the Operational Research Society*. https://doi.org/10.1057/palgrave.jors.2601151

Davis, R. H., D. B. Edelman, A. J. Gammerman. 1992. *Machine-learning algorithms for credit-card applications IMA Journal of Management Mathematics* 4(1) 43-51.

Deakin, E. (1972) *A Discriminant Analysis of Predictors of Business Failure. Journal of Accounting Research*, 10, 167-179. http://dx.doi.org/10.2307/2490225

Durand, D. (1941). *Risk Elements in Consumer Instalment Financing.* National Bureau of Economic Research, New York.

Edward I. Altman 1968. *Financial Ratios, Discriminant Analysis and The Prediction of Corporate Bankruptcy* https://doi.org/10.1111/j.1540-6261.1968.tb00843.x

Equal Credit Opportunity Act (1975), U.S.C., Title 15, Sec 1691 et. seq.

Federal Register (1976). Part 202- Equal Credit Opportunity 41, No. 216 (November 8), 49131-39.

Fisher, E.M (1936) *Linear Discriminant Analysis. Statistics & Discrete Methods of Data Science*, 392, 1-5.

Gombola, Michael J., Mark E. Haskins, J. Edward Ketz and David D. Williams. (1987). *Cash Flow in Bankruptcy Prediction. Financial Management* Vol. 16, No. 4, pp. 55-65 https://doi.org/10.2307/3666109

Hand, D. J., W. E. Henley. 1997. *Statistical classification methods in consumer credit scoring: a review. Journal of the Royal Statistical Society Series*. A 160 523-541.

Hardy, W. E. & Adrian, J. L. (1985). *A linear programming alternative to discriminant analysis in credit scoring. Agribusiness* 1(4) 285-292. http://dx.doi.org/10.1002/1520-6297(198524)1:4<285: AID-AGR2720010406>3.0.CO;2-M

Harold Bierman, Jr., Warren H. Hausman 1970. *The Credit Granting Decision*. https://doi.org/10.1287/mnsc.16.8.B519

Henley, W. E. & Hand, D. J. (1996). *A k-nearest-neighbor classifier for assessing consumer credit risk. Statistician* 65 77-95. http://dx.doi.org/10.2307/2348414

Henley, W. E. 1995. *Statistical Aspects of Credit Scoring*. Ph.D. thesis. Open 83 University, Milton Keynes, U.K.

Henley, W. E. 1995. *Statistical Aspects of Credit Scoring*. Ph.D. thesis. Open 83 University, Milton Keynes, U.K.

James A. Gentry, Paul Newbold and David T. Whitford. 1985. *Classifying Bankrupt Firms with Funds Flow Components. Journal of Accounting Research*. Vol. 23, No. 1 (Spring, 1985), pp. 146-160. https://doi.org/10.2307/2490911

James A. Ohlson. 1980. *Financial Ratios and the Probabilistic Prediction of Bankruptcy. Journal of Accounting Research*. Vol. 18, No. 1 pp. 109-131 https://doi.org/10.2307/2490395

John Banasik, L. C. Thomas 1996. *Does Credit Scoring Make a Difference? The International Review of Retail, Distribution and Consumer Research*

Jos A. M. van Kuelen, Jaap Spronk, A. Wayne Corcoran 1981. *Note-On the Cyert-Davidson-Thompson Doubtful Accounts Model Management Science*, Vol 27, Issue 1 https://doi.org/10.1287/mnsc.27.1.108

Karels, G. and Prakash, A. (1987) *Multivariate Normality and Forecasting of Business Bankruptcy. Journal of Business Finance Accounting*, 14, 573-593. https://doi.org/10.1111/j.1468-5957.1987.tb00113.x

Leonard, K. J. (1993). *Empirical Bayes analysis of the commercial loan evaluation process. Statistics & Probability Letters* 18(4) 289-296. http://dx.doi.org/10.1016/0167-7152(93)90017-D

Lussier, Robert N. (1995). *A nonfinancial business success versus failure prediction model for young firms. Journal of Small Business Management*; Milwaukee Vol. 33, Issue 1: 8.

Makowski, P. (1985). *Credit scoring branches out. Credit World* 75 30-37.

Marc Blum 1974. *Failing Company Discriminant Analysis. Journal of Accounting Research*. Vol. 12, No. 1, pp. 1-25 https://doi.org/10.2307/2490525

Mehta, D. (1968). *The formulation of credit policy models. Management Science* 15(2) 30-50. http://dx.doi.org/10.1287/mnsc.15.2.B30

Michael Lewis, Hopper (1992). *The Self in Self-Conscious Emotions*. https://doi.org/10.1111/j.1540-6834.1992.tb00297.x

Mills, M. (2011). Non-Parametric Methods: The Kaplan-Meier Estimator. *Introducing survival and event history analysis. SAGE Research Methods*, 62-86.

Myers, J. H., E. W. Forgy. 1963. *The development of numerical credit evaluation systems. Journal of the American Statistical Association* 58 799-806.

Narain, B. (1992*). Survival analysis and the credit granting decision.* L. C. Thomas, J. N. Crook, D. B. Edelman, ed. *Credit Scoring and Credit Control*. Oxford

Noel Capon and Deanna Kuhn (1982). *The Journal of Consumer Research.* Vol. 8, No 4 (Mar, 1982). 449-453. https://doi.org/10.1086/208885

Orgler. (1970). *A Credit Scoring Model for Commercial Loans*. *Journal of Money Credit and Banking*, vol.2, issue4, 435-45 http://links.jstor.org/sici?sici=0022-2879%2819701...0.C0%B27&origin=bc

Peto, R., & Pike, M. C. (1973). 352. *Note: Conservatism of the Approximation Σ(O-E)^2/E in the Logrank Test for Survival Data or Tumor Incidence Data*. *Biometrics*, 579-584.

Prinja, S., Gupta, N., & Verma, R. (2010). Censoring in clinical trials: review of survival analysis techniques. *Indian journal of community medicine: official publication of Indian Association of Preventive & Social Medicine*, *35*(2), 217.

R. M. Cyert & H. J. Davidson & G. L. Thompson 1962. *Estimation of the Allowance for Doubtful Accounts by Markov Chains. Management Science*, INFORMS. https://doi.org/10.1287/mnsc.8.3.287

Rich, J. T., Neely, J. G., Paniello, R. C., Voelker, C. C., Nussenbaum, B., & Wang, E. W. (2010). *A practical guide to understanding Kaplan-Meier curves*. *Otolaryngology—Head and Neck Surgery*, *143*(3), 331-336.

Robert A. Eisenbeis (1978). *"A Multivariate Analysis of Industrial Bond Ratings" and the Role of Subordination: A Comment. The Journal of Finance*. Vol. 33, No. 1 pp. 325-335. https://doi.org/10.2307/2326372

Robert O. Edmister 1972. *An Empirical Test of Financial Ratio Analysis for Small Business Failure Prediction. The Journal of Financial and Quantitative Analysis*. https://www.jstor.org/stable/2329929

Rosenberg, E., A. Gleit. 1994. *Quantitative methods in credit management: a survey. Operations Research* 42(4) 589-613.

S. Mcleay A. Omar 2000. *The Sensitivity of Prediction Models to the Non-Normality of Bounded and Unbounded Financial Ratios*. https://doi.org/10.1006/bare.1999.0120

Stalpers, L. J., & Kaplan, E. L. (2018). Edward L. Kaplan and the Kaplan-Meier survival curve. *BSHM Bulletin: Journal of the British Society for the History of Mathematics*, *33*(2), 109-135.

Stepanova, M., L. C. Thomas. 2001. *PHAB scores: Proportional Hazards Analysis Behavioral Scores. Journal of the Operational Research Society* 52(9) 1007- 1016.

Thomas, L. (2002) and Stepanova, M. *Survival Analysis Methods for Personal Loan Data. Operations Research*, 50, 277-289. https://doi.org/10.1278/opre.50.2.277.426

Troy Segal, Janet Berry-Johnson, Skylar Clarine https://www.investopedia.com/terms/f/five-c-credit.asp

Tsai, K., Pollock, K. H., & Brownie, C. (1999). Effects of violation of assumptions for survival analysis methods in radiotelemetry studies. *The Journal of wildlife management*, 1369-1375.

Venkat Srinivasan & Yong H. Kim 1987. *Credit Granting: A Comparative Analysis of Classification Procedures. The Journal of Finance*, Vol 42, Issue 3, p. 665-681. https://doi.org/10.1111/j.1540-6261.1987.tb04576.x

Wang, E. W. (2010). Rich, J. T., Neely, J. G., Paniello, R. C., Voelker, C. C., Nussenbaum, B., A practical guide to understanding Kaplan-Meier curves. *Otolaryngology—Head and Neck Surgery*, *143*(3), 331-336.

Wiginton, J.C. (1980). *A Note on the Comparison of Logit and Discriminant Models of Consumer Credit Behavior. The Journal of Financial Quantitative Analysis,* 15, 757-770. https://doi.org/10.2307/2330408

Zavgren, Christine. (1983). *"The prediction of corporate failure: the state of the art." Journal of Accounting Literature* 2.1: 1-38.

Zmijewski M. E. 1984. *Methodological Issues Related to the Estimation of Financial Distress Prediction Models. Journal of Accounting Research*. https://doi.org/10.2307/2490859