UNIVERSITY OF NAIROBI

SCHOOL OF MATHEMATICS

A research project submitted to the school of mathematics in partial fulfillment of the requirements for the award of the degree of Master of Science in social statistics of the University of Nairobi.

2016.

BICHANG'A SIMION NYAKORA

MSC SOCIAL STATISTICS

I56/74586/2014

Modeling teacher attrition using Univariate frailty model.

DECLARATION

This project is my original work and has not been presented for a degree in any other university

…………………………………. ………………………………

Signature Date

This project has been submitted for examination with my approval as University supervisor

…………………………………. ………………………………

Signature Date

DEDICATION

I dedicate this project to my beloved wife Lucinah Bosibori and my dear children June Ray and Jayd Ram for their prayers and infinite support all through my coursework.

ACKNOWLEDGEMENTS

I'm grateful to the almighty God for the gift of life and good health all through.

I would like to thank my able supervisor Dr. Nelson Owuor for his positive and charismatic guidance and support.

I also thank my colleagues for constant guidance and assistance especially Victor Ouma for his tireless support guidance and inspiration.

Lastly, thanks to everybody who prayed, encouraged and challenged me.

Thanks you all.

ABSTRACT

Researchers and policy developers have done a lot of studies on the reasons fueling teacher attrition in secondary schools. I contribute to these findings by evaluating the influence of gender, marital status, teaching subjects, age, education level and place of residence of a teacher on the teacher's decision to transfer out of a sub county. Univariate frailty model with covariates was used to model teacher attrition using information from simulated data. The study objectives were to; develop a frailty model for teacher attrition, evaluate the factors influencing teacher attrition using frailty model and to determine the trend on teacher attrition rate. All the factors under investigation were found to have a significant influence on teacher transfer decision except the place of residence. The study as well found out that there were other unobserved factors which influence a teacher's decision to transfer out of a given region. The data used in this study was randomly generated. Information investigation was done utilizing R-3.2.5 rendition.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

Cdf Cumulative density function

Pdf Probability density function

Mgf moment generating function

ICTs Information communication technologies

TSC Teachers service commission

AIC Akaike information criteria

BIC Bayesian information criteria

UNESCO United Nations education, science and cultural organization

KNUT Kenya national union of Teachers

SYMBOLS

E expectation

f generic symbol for a probability density function

F generic symbol for a cumulative density function

Γ Gamma function

L Laplace transform

λ generic symbol for a (baseline) hazard function

P(A) probability of event A

S generic symbol for a survival function

V variance

$X \sim Exp(\lambda)$ exponential distributed random variable with parameter λ

$X \sim \text{Weibull}(a, b)$ Weibull distributed random variable with parameters a, b

$X \sim \Gamma(k, \lambda)$ Gamma distributed random variable with parameters k, λ

τ Kendall's tau

CHAPTER 1

1.0 Introduction

1.1 Background

There is bottomless writing which has made an endeavor to diagnose the state of high teacher turnover in several developed nations like Herbert and Ramsay (2004) in United States, Finlayson (2003) in Scotland and Santiago (2001) in Britain with a portion of the studies reporting teacher attrition having reached national crisis. The teacher turnover situation in Sweden, Germany, New Zealand and some states in USA has been worsening.

In developing nations, the issue of educator turnover is relatively serious. A study by Xaba, (2003) in South and some countries in Africa like Zambia and Malawi pointed out that the issue of teacher attrition had almost reached a calamitous stage. Most African countries are faced with serious teacher deficiency due to high attrition rates especially in remote rural schools that cannot attract qualified instructors, a part from the head teacher. As a result, untrained teachers are over 33% of the schools' staff. In Malawi according to Kadzamira (2005), there was a shocking urban bias in the distribution of educational resources. Low job satisfaction makes it difficult to properly staff rural schools. Remote rural schools are chronically understaffed due to high instructor turnover and the refusal of educators to be conveyed to schools in these areas. It is hard to staff remote institutions, once a teacher is lost; it is very difficult to find a replacement. These researchers all agree that there is a great problem of teacher turnover in developing countries, especially in rural remote areas.

The Kenyan government and UNESCO (2012) end of Decade assessment of instruction in Kenya bemoaned that in spite of the fact that the understudy educator proportion at the national level may demonstrate that the nation has accomplished the suggested proportion of 45:1, there are still local inconsistencies in the Coast and North Eastern areas, where the student instructor proportion can be as high as 63:1. Regardless of the endeavors of the administration to recruit educators, the instructor deficiency still continues. It is doubtful that the loss of qualified educators from the calling for any reason influences Kenya's monetary improvement, especially in the exploratory, mechanical, and professional divisions, and target which the Government of Kenya is endeavoring to accomplish through education. It is against this foundation that this study set out to research the status of the teacher turnover and inspect a portion of the explanatory variables for the educator turnover in public high schools.

There is developing proof that have tried to investigate the difficulties defying teacher education, likely variables to explain the high teacher turnover as well as improvements of instructor training in Africa (Ondaro,2004)A portion of the difficulties identify with: the increasing gap between the interest and supply of educators particularly in science and arithmetic; the expanding interest for better quality instructors and instructor education; the requirement for social and expert control in connection to quality confirmation; The difficulties of overseas training; improvements in Information Communication Technologies (ICTs) and the resultant requirement for utilizing present day ICTs in the preparation of teachers; weight for national aggressiveness in a globalized information based economy against a current supply of untrained and under prepared instructors in numerous African nations; and, the failure of the customary private college model of educator instruction and preparing to satisfactorily meet either present or anticipated interest for educator training . Among these difficulties, instructors' turnover has been given minimal consideration in the nation. Teaching is still the poorest paying employments in the nation, quality instructors cannot be held under the present terms and states of administration.

As indicated by Oplatka (2007) teaching was among the respected jobs for Africans before independence; it was a noble profession then in Kenya. A teacher was regarded highly as a source of knowledge therefore commanded respect, after freedom, teaching lost its glory and could not attract most gifted students. Factors influencing teacher attrition include gender, professional qualifications, job satisfaction, teaching experience, social-cultural factors, and career commitment. Educators stay on the job if physical, economic, social status and security needs related to their working conditions as well as adequate provision of salary, good supervision, teaching materials, small classes, timely and premium pay improve job satisfaction. These elements if not present in a working place brings down the assurance of the instructors, prompting the turnover.

1.2 Statement of the Problem

The Government of Kenya is aware of the magnitude of the issue of high teacher turnover, however just expresses that the present interest driven employment arrangement was set up to address the uneven distribution of educators and instructor deficiencies (Government of Kenya and UNESCO, 2012). Because of the stop on new employment, the commission has just been replacing teachers who leave the profession through death. Reality on the ground is that there is an intense lack of educators to cater for higher demand after the introduction of essential and free day secondary education, where the administration accommodates free educational cost charges among different components. In 2010, with an aim of conquering lack of teachers, the government employed 18,060 instructors on contract terms as a makeshift measure to ease the instructor deficiency. There is an issue of work turnover among the educators in public secondary schools in Kenya, particularly those working in remote areas.

In spite of the fact that the Kenya government urges educators to stay in hardship regions by giving incidental advantages, most instructors working in such zones feel that what is offered is too little contrasted with what their partners in urban zones make as an afterthought by participating in businesses. These elements could prompt absence of dedication for instructors working in rural regions as is obvious in their unsteadiness in the teaching profession and low confidence in teaching and thus unsuitable execution of day by day school duties and activities. Instructors have been leaving teaching for better paying jobs despite the hardship stipend by the Teachers' Service Commission (TSC).

Generally, In Kenya, the number of teachers leaving the profession has tremendously increased in recent times. Likewise the rate of teacher transfer from specific areas has been generally high. It is not clear the reasons precipitating the high turnover of teacher and therefore the study seeks to evaluate how well certain factors contribute to the teacher attrition. Many studies have been using regression methods in determining the influence of factors on the teacher attrition but the risk of factors are rarely captured hence as they are not easily seen this study seek to include the frailty in modeling the risks of teacher attrition.

1.3 Main Objectives

To model teacher attrition using information from simulated data.

1.4 Specific objectives

1. Develop a frailty model for teacher attrition.
2. Evaluate factors influencing teacher attrition using frailty model.
3. Determine the trend on teacher attrition rate.

1.5 Justification /significance of the study

This study is of extraordinary significance as it can assist the administration and other education partners with knowing the state of teacher transfer and reasons they transfer from a given place and in extension the whole of Kenya. This information can help the government to know whether the teacher recruitment being done annually is sufficient in replacing the teachers who transfer, are fired, who die or dessert duty. The government as well can know the reasons for teacher attrition and address them to improve the retention of teachers in the profession.

1.6 Scope

The data used in this thesis was generated randomly using uniform distribution because of the inability to collect actual data due to time constraint. The data was analyzed using univariate frailty model with covariates to with the main aim of evaluating attrition (transfer) rate of secondary school teachers outside a school sub county.

1.7 Limitations

The main limitation was the inability to get actual data on teacher transfer due to bureaucracy and time constraints.

CHAPTER 2

2.0 Literature review

Linda Kuzan, M. (1988) investigated the factors influencing tutor attrition in Wisconsin's education system. Alpha factoring and Principal components analysis helped in reducing the variables from thirty seven to thirteen. On the basis of the stated factors, discriminant analysis came in handy in determining if precise predictions could be reached whether a teacher left or remained in teaching. The attrition rate in both general and special education was calculated. In general, the findings indicated significant higher attrition rates for special education teachers, teacher on emergency certification, teachers at secondary level, music, teacher with undergraduate training, foreign language teachers and teachers with less than six years experience in teaching.

John, Y. et al (2008), did a study in five vast metropolitan territories in upstate New York. The study concentrated on instructor's choice to leave a school locale or to leave instructing utilizing the prentice-Gloeckler Meyer method for unmeasured heterogeneity. The study included educator qualities like pay rates, training level, grade level, subjects taught and confirmation. These characteristics were found to contribute to teacher's decision to quit teaching or transfer .The strength of this technique is that it uses random effects to incorporate unobserved heterogeneity. In conclusion, the study pointed out that the probability of a teacher transfer or quitting can be reduced by considerable salary increase.

Ghadimi et al, (2011) studied family history of the growth on the survival of the patients with gastrointestinal cancer in northern parts of Iran, using frailty models with an objective of determining the most common risk factors affecting the chance of surviving of the ailing people with GI tract cancer. Their approach was through a comparison of parametric models (Weibull, Exponential, Log-normal, and the Log-logistic) with frailty. The risk factors studied included; family history of cancer, age, sex, marital status, smoking status, engagement, tribe, state of taking medicine, education level, place of residence, cancer type, and migration status. Hazard ratio was used to interpret the death risk for the two models (fitted with and without frailty) and Akaike information criterion (AIC) compared to select the best model. The parametric models were compared, the log-logistic model with gamma frailty was better than the others and using that model, gender and the family history of the cancer were found to be significant predictors.

Hassan Aslami (2013), studied why secondary school teachers leave the profession in Afghanistan. The study used mixed approach in collection and analysis of data. The questionnaire captured information such as school type, gender, age, teaching subjects and years of service of teachers. Further, the study indicated that apart from salary, other factors influencing teacher attrition include; marital status, safety, low social status, school distance, ghost teachers and unfair transfers

Sharon, J. et al, (2014) did a study on Institutional factors contributing to teacher attrition in public schools as well as examining the relationship between teachers' individual factors and tutor turnover in Baringo . Qualitative data was analyzed with content analysis while quantitative data were analyzed using descriptive statistics, i.e. the mode, median, and the mean. In addition, the Pearson Moment correlation coefficient was used to ascertain the relationship between the independent factors and the dependent factor. The study considered a wide range of factors for accurate picture on factors influencing teacher turnover. The study pointed that the key factors influencing teacher attrition were present assignment, level of education, marital status and institutional factors like school category was also found to be an influential contributor to teacher attrition.

Khuda, B. M. (2016) evaluated the determinants of teacher attrition at secondary school level in Punjab with the aim of investigating factors influencing teacher attrition. Multiple regression was used in analyzing the data and the findings revealed that apart from the school climate, salary, principal behavior ,work load and contractual employment, there are other factors influencing teacher attrition that were never studied since joint influence of the five determinants accounted for 65% of the attrition.

.

CHAPTER 3

3.0 Methodology

3.1 Data simulation

The data used in the study was generated for the variables of interest. The variables are as listed in the table below;

Table 1: variables of interest

No.	Variable	Levels	Data type	Factor or integer
1	Sub-County	10	Numeric	Integer
2	Teacher	3000	Numeric	Integer
3	Gender	0=Male,1=Female	Categorical	Factor
4	Residence	0=Lari, 1=Nairobi, 2=Kikuyu,3=Limuru	Categorical	Factor
5	Education levels	0=diploma, 1=Bed, 2=MSc, 3=PhD	Categorical	Factor
6	Teaching subject	0=Home_science, 1=Mathematics, 2=English, 3=Kiswahili, 4=CRE, 5=Geography, 6=Agriculture, 7=Business_studies, 8=History, 9=Physics,10=Biology, 11=Chemistry, 12=Computer _studies	Categorical	Factor
7	Present Assignment	0= HoD,1= Principal, 2=Deputy_head, 3= Teacher	Categorical	Factor
8	Age	Open	Numeric	Integer
9	Marital status	0=Single, 1=Married,	Categorical	Factor

		2=Divorced		
10	School category	0=National,1=Extra-County, 3=County, 4=CDF	Categorical	Factor
11	Duration of stay	0-20	Numeric (time to event)	Integer
12	Reason for leaving	0=Other (death, sacking, resignation, retirement.) 1=Transfer	String	Censored

3.1.1 Frailty model

The frailty model for the study is a Univariate frailty model with covariates as shown

$$\mu(t, Z, X) = Z\mu_0(t)\exp(\beta^T X) \quad [1]$$

Where;

μ_0 (t) - Is the baseline hazard function,

β - Is the vector of regression coefficients for the covariates

X - Is the vector of observed covariates

Z -is the frailty variable. (The frailty Z is a random variable changing over the population that reduces (Z<1) or adds (Z>1) the individual risk)

3.1.2 Assumptions for the model

Survival times should be positively associated

Homogeneity of the population

$\mu_0(t)$ Is the baseline hazard. Under the parametric approach, the baseline hazard is a parametric function and the vector of its parameters, say ψ, is estimated in line with the coefficients of regression and the frailty (risk) parameter(s). Here the Weibull distribution of the hazard function is adopted.

3.1.3 Distribution for the frailty

The frailty is gamma distributed with a random variable $x \sim \Gamma(\gamma, \beta)$ with a standard probability density function given by

$$f(x) = \frac{x^{\gamma\text{-}1} \exp(-x)}{\Gamma(\gamma)}, \qquad x \geq 0, \gamma > 0 \qquad [2]$$

Where γ is the shape parameter while β is scale parameter. Γ is the gamma function which has the formula, $\Gamma(a) = \int_0^\infty t^{\alpha-1} \ell^{-t} dt$. It corresponds to a gamma distribution $\Gamma(\gamma, \beta)$ with β fixed to 1 for identifiability.

The hazard function of the gamma distribution is

$$h(x) = \frac{x^{\gamma-1} \ell^{-x}}{\Gamma(\gamma) - \Gamma_x(\gamma)} \qquad \mathrm{x} \geq 0, \gamma > 0 \qquad [3]$$

Where $\Gamma_x(\gamma)$ is the incomplete gamma function given by, $\Gamma_x(a) = \int_0^{\mathrm{x}} \mathrm{t}^{\alpha\text{-}1} \ell^{-t} dt$ while the cumulative hazard function is given by;

$$H(x) = -\log\left(1 - \frac{\Gamma_x(\gamma)}{\Gamma(\gamma)}\right) \qquad \mathrm{x} \geq 0, \gamma > 0 \qquad [4]$$

The survival function, $S(t) = 1 - \frac{\Gamma_x(\gamma)}{\Gamma(\gamma)} \qquad \mathrm{x} \geq 0, \gamma > 0 \qquad [5]$

For the gamma distribution, the Kendall's tau, which measures the association between two event times from the same group in the multivariate case, can be computed as

$$\tau = \frac{\theta}{\theta + 2} \in (0,1) \qquad [6]$$

Where θ is the variance of the frailty term.

3.1.4 Distribution of the baseline hazard

In this study the distribution of the baseline hazard is weibull. The univariate frailty model is a generalization of the exponential model with non-negative parameters. The weibull model was chosen because of its great flexibility and the different forms of its hazard function which makes it convenient to model empirical work and then again from the simplicity of the hazard and survival function. The equation for the standard two-parameter weibull distribution with $\alpha = 1$ is;

$$f(x) = \gamma\, x^{(\gamma-1)} \ell^{\left(-x^{\gamma}\right)} \qquad x \geq 0, \gamma > 0 \qquad [7]$$

Where γ is the shape parameter while α is the scale parameter.

The hazard function, h(x) of the Weibull distribution, cumulative hazard function, H(x), and the survival function, S(x) are;

$$h(x) = \gamma\, \mathrm{x}^{(\gamma\text{-}1)} \qquad x \geq 0, \gamma > 0 \qquad [8]$$

$$H(x) = \mathrm{x}^{\gamma} \qquad x \geq 0, \gamma > 0 \qquad [9]$$

$$S(x) = \exp\text{-}\left(\mathrm{x}^{\gamma}\right) \qquad x \geq 0, \gamma > 0 \qquad [10]$$

3.2 Likelihood approach to frailty model

3.2.1 Survival likelihood

Survival data consists of event times and censored observations under random censoring. The likelihood function for survival data is given by

$$L = \prod_{j=1}^{n} \left[(1 - Gj(t)) f_j(t)\right]^{\delta_j} \left[(1 - F_j(t)) \mathrm{g}_{\mathrm{j}}(t)\right]^{1-\delta_j}, \qquad [11]$$

Where the censoring indicator is δ_j, g is the density function and G is the cumulative distribution function of the censoring time, f and F are the density function and the cumulative distribution function of the event time, respectively.
We can ignore the censoring time distribution in the likelihood function since it is independent of the required parameters of the survival function. With right censoring, the likelihood function for the j_{th} subject is ;

$$L = \prod_{j=1}^{n} \left(f_j(t)\right)^{\delta_j} \left(S_j(t)\right)^{1-\delta_j} \qquad [12]$$

Following the idea above, the likelihood function for the j_{th} subject in the i_{th} subgroup is given by

$$L_i = \prod_{j=1}^{ni} \left(f_{ij}(t)\right)^{\delta_{ij}} \left(S_{ij}(t)\right)^{1-\delta_{ij}} \qquad [13]$$

Since $h_{ij}(t) = \dfrac{f_{ij}(t)}{S_{ij}(t)}$,then $f_{ij} = h_{ij}(t)\ S_{ij}(t)t)$. We can rewrite the conditional likelihood function in the form of

$$L_i = \prod_{j=1}^{n_i} \left(h_{ij}(t)\right)^{\delta_{ij}}\ \mathrm{S}_{\mathrm{ij}}(t) \qquad [14]$$

Following these ideas, we can easily derive the forms of the conditional and marginal likelihood functions of the frailty model.

Cox proportional hazards model for frailties is given by

$$h_{ij}(t) = h_0(t)u_i\, \ell^{\beta^t Z_{ij}} \qquad [15]$$

Where $u_i's$ are independent and identically distributed random sample from a distribution with mean of 1 and some unknown variance of θ. The equation (15) be written as,

$$\frac{f(t_{ij})}{S(t_{ij})} = \mathrm{h}_0(t)\mathrm{u}_{\mathrm{i}}\ell^{\beta^t Z_{ij}}\ , \qquad [16]$$

Integrating both sides of the Eq. (16), we can get the expression for the survival function.

$$-l\mathrm{n}\left(S_{ij}(t)\right) = H_0(t)u_i\ell^{\beta^t Z_{ij}} \qquad [17]$$

Therefore,

$$S_{ij}(t) = \exp\left(-H_0(t)u_i\ell^{\beta^{\prime} Z_{ij}}\right) \qquad [18]$$

For the i_{th} subgroup, its conditional likelihood function is ;

$$L_i(\psi,\beta \mid u_i)=\prod_{j=1}^{n_i}(h_0(t)u_i\ell^{\beta^t Z_{ij}})^{\delta_{ij}}\ \ell^{-H_0(t)u_i\ \ell^{\beta^t Z_{ij}}} \qquad [19]$$

Where ψ is the baseline hazard's vector of parameters i.e. $\psi=(\lambda,\rho)$. Hence, the marginal likelihood function for the i_{th} subgroup is

$$L_i(\psi,\theta,\beta)=\prod_{j=1}^{n_i}\int_0^{\infty}(h_0(t)u\ell^{\beta^t Z_{ij}})^{\delta_{ij}}\ \ell^{-H_0(t)u\ \ell^{\beta^t Z_{ij}}}\ g(u)d(u) \qquad [20]$$

Where *g* (*u*) is the probability distribution function of frailties $u_{1,\dots\dots,}u_G$.

To obtain the marginal log likelihood for the gamma frailty model, let u_1 be identical and independently distributed (iid) sample of gamma random variables with density function

$$g(u)=\frac{u^{1/\theta-1}\ell^{-u/\theta}}{\Gamma(1/\theta)\theta^{1/\theta}},\mathrm{u}>0,\theta>0, \qquad [21]$$

With *E* (*U*) = 1 and *Var* (*U*) = *θ*. Larger values of *θ* indicate that there is a higher degree of heterogeneity among groups and strong association within groups. First, we show that gamma frailties can be integrated out in the conditional survival likelihood. This would lead to explicit and simple marginal likelihood function which only contains the parameters of interest. The marginal likelihood function for the i_{th} group is given by

$$L_i(\psi,\theta,\beta)=\prod_{j=1}^{n_i}\int_0^{\infty}(h_0(t)u\ell^{\beta^t Z_{ij}})^{\delta_{ij}}\ \ell^{-H_0(t)u\ \ell^{\beta^t Z_{ij}}}\ \frac{u^{1/\theta-1}\ell^{-u/\theta}}{\Gamma(1/\theta)\theta^{1/\theta}}d(u) \qquad [22]$$

Rearranging the terms in Eq. (22), we obtain the following expression

$$L_i(\psi,\theta,\beta)=\prod_{j=1}^{n}h_0(t)^{\delta_{ij}}\ \ell^{\beta^t Z_{ij}\delta_{ij}}\int_0^{\infty}\frac{u^{1/\theta+d_i-1}\ell^{-u/\theta}\exp\left(-\sum_{j=1}^{n_i}H_0(t)u\ell^{\beta^t Z_{ij}}\right)}{\Gamma(1/\theta)\theta^{1/\theta}}du,$$

$$=\prod_{j=1}^{n_i}h_0(t)^{\delta_{ij}}\ \ell^{\beta^t Z_{ij}\delta_{ij}\frac{\Gamma(1/\theta+d_i)\theta^{(1/\theta+d_i)}}{\Gamma(1/\theta)\theta^{1/\theta}}\int_0^{\infty}\frac{u^{1/\theta+d_i-1}\exp\left(-u\left(1/\theta+\sum_{j=1}^{n_i}H_0(t)\ell^{\beta^t Z_{ij}}\right)\right)}{\Gamma(1/\theta+d_i)\theta^{1/\theta+d_i}}}du, \qquad [23]$$

where $d_i = \sum_{j=1}^{n_i} \delta_{ij}$

We integrate out the frailty term u. The term under the integral is the moment generating function (mgf) of a gamma distribution with a pdf $\Gamma\left(\frac{1}{\theta + d_i}, \frac{1}{\theta}\right)$ using this fact; we can derive the expression for marginal likelihood function as

$$L_i(\psi, \theta, \beta) = \prod_{j=1}^{n_i} h_0(t)^{\delta_{ij}} \ell^{\beta^t Z_{ij} \delta_{ij}} \frac{\Gamma\left(\frac{1}{\theta} + d_i\right) \theta^{\left(\frac{1}{\theta} + d_i\right)}}{\Gamma\left(\frac{1}{\theta}\right) \theta^{\frac{1}{\theta}} \theta^{\left(\frac{1}{\theta} + d_i\right)} \left(\frac{1}{\theta} + \sum_{j=1}^{n_i} H_0(t) \ell^{\beta^t Z_{ij}}\right)^{\frac{1}{\theta} + d_i}}$$

$$\int_0^\infty \frac{u^{\frac{1}{\theta} + d_i - 1} \ell^{-u\left(\frac{1}{\theta} + \sum_{j=1}^{n_i} H_0(t) \ell^{\beta^t Z_{ij}}\right)} \left[\frac{1}{\theta} + \sum_{j=1}^{n_i} H_0(t) \ell^{\beta^t Z_{ij}}\right]^{\left(\frac{1}{\theta} + d_i\right)}}{\Gamma\left(\frac{1}{\theta} + d_i\right)} du$$

$$= \prod_{j=1}^{n_i} h_0(t)^{\delta_{ij}} \ell^{\beta^t Z_{ij} \delta_{ij}} \frac{\Gamma\left(\frac{1}{\theta} + d_i\right)}{\Gamma\left(\frac{1}{\theta}\right) \theta^{\frac{1}{\theta}} \left(\frac{1}{\theta} + \sum_{j=1}^{n_i} H_0(t) \ell^{\beta^t Z_{ij}}\right)^{\frac{1}{\theta} + d_i}} \qquad [24]$$

$$\int_0^\infty \frac{u^{\frac{1}{\theta} + d_i - 1} \ell^{-u\left(\frac{1}{\theta} + \sum_{j=1}^{n_i} H_0(t) \ell^{\beta^t Z_{ij}}\right)} \left[\frac{1}{\theta} + \sum_{j=1}^{n_i} H_0(t) \ell^{\beta^t Z_{ij}}\right]^{\frac{1}{\theta} + d_i}}{\Gamma\left(\frac{1}{\theta} + d_i\right)} du$$

We can see that the term under the integral is the probability distribution function (pdf) of

$\Gamma\left(\frac{1}{\theta} + d_i, \frac{1}{\theta} + \sum_{j=1}^{n_i} H_0(t) \ell^{\beta^t Z_{ij}}\right)$, which integrates to 1

The final marginal likelihood function is

$$L_i(\psi,\theta,\beta)=\frac{\Gamma\left(\frac{1}{\theta}+d_i\right)\prod_{j=1}^{n_i} h_0(t)^{\delta_{ij}}\,\ell^{\beta^t Z_{ij}\delta_{ij}}}{\left(\frac{1}{\theta}+\sum_{j=1}^{n_i} H_0(t)\ell^{\beta^t Z_{ij}}\right)^{\left(\frac{1}{\theta}+d_i\right)}\Gamma\left(\frac{1}{\theta}\right)\theta^{\frac{1}{\theta}}} \qquad \textbf{[25]}$$

Taking the logarithm of this expression and summing over the G clusters, we obtain the marginal likelihood function, $l(\psi,\theta,\beta)$

$$l(\psi,\theta,\beta)=\sum_{i=1}^{G}\left[d_i\log(\theta)-\log\left(\Gamma\left(\frac{1}{\theta}\right)\right)+\log\left(\Gamma\left(\frac{1}{\theta}+d_i\right)-\left(\frac{1}{\theta}+d_i\right)\log\left(1+\theta\sum_{j=1}^{n_i}H_0(t)\ell^{\beta^t Z_{ij}}\right)+\right)\right]$$

$$\sum_{j=1}^{n_i}\delta_{ij}\left(\beta^t Z_{ij}+\log(h_0(t))\right)\Bigg)\Bigg] \qquad [26]$$

Considering a gamma frailty with weibull baseline hazard rate, marginal log likelihood function is

$$l(\psi,\theta,\beta)=\sum_{j=1}^{G}\left[d_i\log(\theta)-\log\left(\Gamma\left(\frac{1}{\theta}\right)\right)+\log\left(\Gamma\left(\frac{1}{\theta}+d_i\right)\right)\text{-}\left(\frac{1}{\theta}+d_i\right)\log\left(1+\theta\sum_{j=1}^{n_i}\left(\frac{t}{\alpha}\right)^{\eta}\ell^{\beta^t Z_{ij}}\right)+\right.$$

$$\left.\sum_{j=1}^{n_i}\delta_{ij}\left(\beta^t Z_{ij}+\log\left(\frac{\eta\,\mathrm{t}^{\eta\text{-}1}}{\alpha^{\eta}}\right)\right)\right] \qquad [27]$$

Taking the first derivative of the gamma frailty model with weibull baseline hazard and one covariate

$$\frac{\partial l(\eta,\alpha,\theta,\beta)}{\partial\alpha}=\sum_{i=1}^{G}\left[\frac{(1+\theta\,d_i)\frac{n}{\alpha}\sum_{j=1}^{n_i}\left(\frac{t}{\alpha}\right)\eta\ell^{\beta^t Z_{ij}}Z_{ij}}{\left(1+\theta\sum_{j=1}^{n_i}\left(\frac{t}{\alpha}\right)\eta\,\ell^{\beta^t Z_{ij}}\right)}-\mathrm{d_i}\,\frac{\eta}{\alpha}\right] \qquad \textbf{[28]}$$

$$\frac{\partial l(\eta,\alpha,\theta,\beta)}{\partial \eta}=\sum_{i=1}^{G}\left[\frac{-(1+\theta d_i)\sum_{j=1}^{n_i}\left(\frac{t}{\alpha}\right)\eta\ell^{\beta^t Z_{ij}}\log(t)}{\left(1+\theta\sum_{j=1}^{n_i}\frac{t}{\alpha}\eta\ell^{\beta^t Z_{ij}}\right)}+\sum_{j=1}^{n_i}\delta_{ij}\left(\frac{1}{\eta}+\log(t)\right)\right] \quad [29]$$

$$\frac{\partial l(\eta,\alpha,\theta,\beta)}{\partial \theta}=\sum_{i=1}^{G}\left[\frac{-\left(\frac{1}{\theta}+d_i\right)\sum_{j=1}^{n}\left(\frac{t}{\alpha}\right)\eta\ell^{\beta^t Z_{ij}}}{1+\theta\sum_{j=1}^{n_i}\left(\frac{t}{\alpha}\right)\eta\ell^{\beta^t Z_{ij}}}+\frac{1}{\theta}\log\left(1+\theta\sum_{j=1}^{n_i}\frac{t}{\alpha}\eta\ell^{\beta^t Z_{ij}}\right)+\frac{d_i}{\theta}+\frac{\Gamma'\left(\frac{1}{\theta}\right)}{\theta^2\Gamma\left(\frac{1}{\theta}\right)}\right] \quad [30]$$

$$\frac{\partial l(\eta,\alpha,\theta,\beta)}{\partial \beta}=\sum_{i=1}^{G}\frac{-(1+\theta d_i)\sum_{j=1}^{n_i}\left(\frac{t}{\alpha}\right)\eta\ell^{\beta^t Z_{ij}}Z_{ij}}{\left(1+\theta\sum_{j=1}^{n_i}\left(\frac{t}{\alpha}\right)\eta\,\ell^{\beta^t Z_{ij}}\right)}+\sum_{j=1}^{n_i}\delta_{ij}Z_{ij} \quad [31]$$

To obtain the maximum likelihood estimates equate the first order derivatives to 0 and solving for the required parameters. Since the equation is nonlinear, it can only be solved using algorithms such as Newton Raphson.

3.3 Model Selection

To compare the covariates that explain teacher attrition, 93 models were developed with varying covariates. I applied the information based criteria to select the best model. AIC is given by the expression;

$AIC = \quad 2log\ (L) + 2k$,

Where L is the maximized likelihood value and k is the number of parameters in the model. BIC is given by the expression

$BIC = \quad 2log\ (L) + kln\ (N)$,

where N is the total sample size.

The model with the smallest AIC, BIC value is considered a better fit. Model 91 with AIC value of 5682.665 and BIC value of 5844.837 was chosen to be the best model.

CHAPTER 4

4.0 Data Analysis and Results

4.1 Exploratory analysis of the data

Box plot

Figure 1: Box Plot (Age, duration stay)

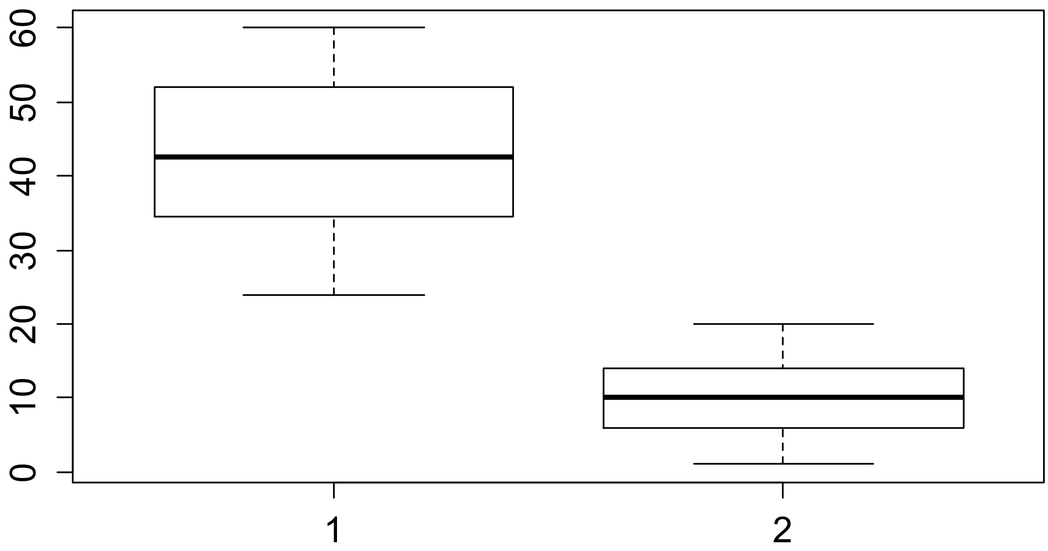
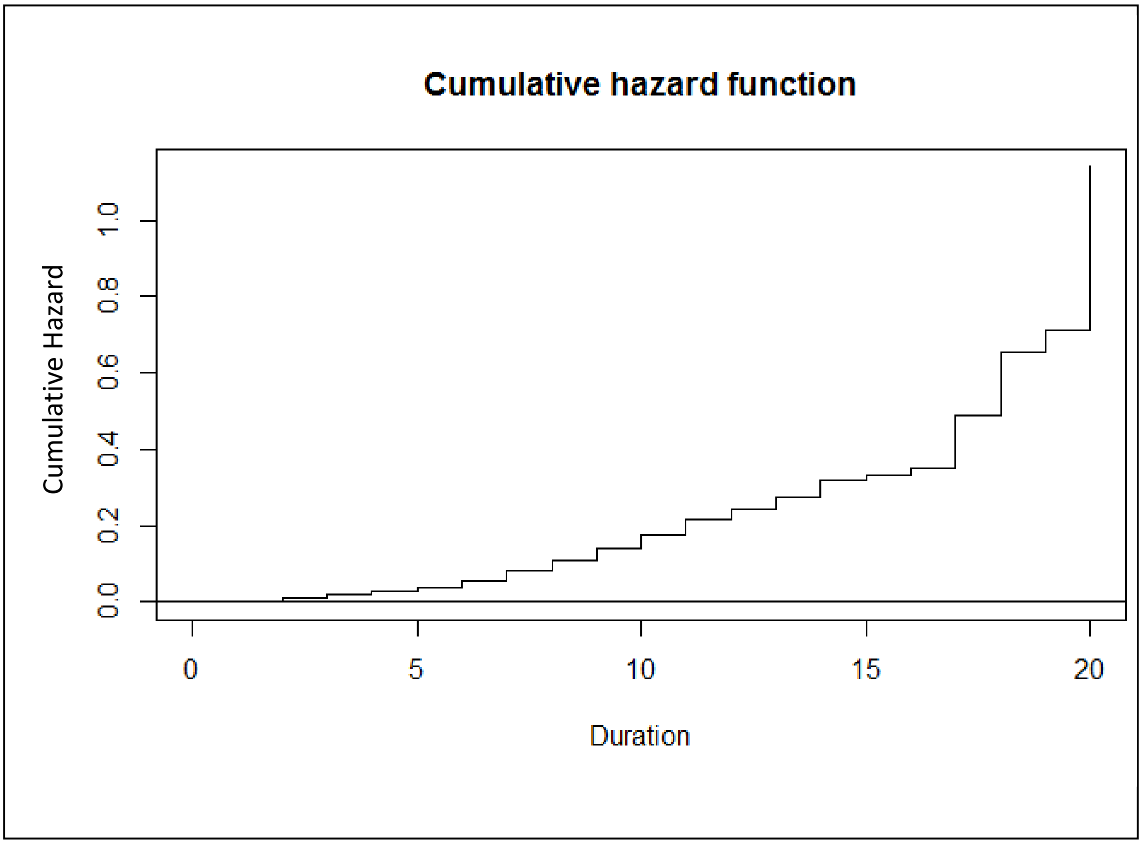


Figure 2: Cumulative Hazard Plot



The cumulative hazard (H(t)) is the probability of failure at time t given survival up to time t. The cumulative hazard function indicates that the transfer rate increases with time, by the end of 20 years all teachers will have transferred.

4.2 Univariate Frailty model analysis

4.2.1 Significant model

Both AIC and BIC selected model 91 as the best model. (attrition.mod91)

Frailty distribution: gamma

Baseline hazard distribution: Weibull

Log likelihood: -2814.332

Table 2: Best Model/Model 91

Coefficients	**Estimate**	**Se**	**p-value**	
theta	0.22	0.105		
rho	2.439	0.075		
lambda	0	0		
Covariates	Estimates	se	p-value	
Gender male	0.458	0.093	0	***
ResidenceLari	-0.25	0.152	0.099	.
ResidenceLimuru	0.258	0.124	0.037	*
ResidenceNairobi	0.211	0.101	0.038	*
Educationlevel Diploma	1.053	0.109	0	***
EducationlevelMSc	0.027	0.115	0.816	
EducationlevelPhD	-0.045	0.143	0.753	
Age	-0.032	0.004	0	***
MaritalstatusSingle	0.392	0.087	0	***
TeachingsubjectsBiology	1.458	0.276	0	***
TeachingsubjectsBusiness_studies	1.081	0.271	0	***
TeachingsubjectsChemistry	0.239	0.268	0.372	
TeachingsubjectsComputer_Studies	1.136	0.221	0	***
TeachingsubjectsCRE	1.754	0.241	0	***
TeachingsubjectsEnglish	0.607	0.23	0.008	**
TeachingsubjectsGeography	0.946	0.232	0	***
TeachingsubjectsHistory	-0.378	0.298	0.205	
TeachingsubjectsHome_science	0.575	0.39	0.141	
TeachingsubjectsKiswahili	0.458	0.249	0.066	.
TeachingsubjectsMathematics	0.745	0.227	0.001	**
TeachingsubjectsPhysics	0.694	0.232	0.003	**
PresentassignmentHoD	0.627	0.14	0	***

PresentassignmentPrincipal	0.061	0.111	0.584	
PresentassignmentTeacher	-0.149	0.13	0.252	

Signif. Codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1   Kendall's Tau: 0.099

4.3 Coefficients

When theta is significantly different from zero, it means that there exists heterogeneity between key subjects that is explained by non-observed covariates. Hypothesis of interest is

$$H_0 : \theta = 0$$
$$\mathrm{H}_1 : \theta \neq 0$$

$$\mathrm{Z} = \frac{\theta}{\mathrm{se}(\theta)} = \frac{0.22}{0.105} = 2.095$$

$$Z_{\left(\frac{\alpha}{2}\right)} = 1.96$$

Since 2.095>1.96, we reject the null hypothesis and conclude that theta is significantly different from zero.

The positive value of the shape parameter, Rho =2.439, in the joint model is indicating that the transfer rate is positively associated with time.

4.4 Covariates

4.4.1 Gender

The estimate of male is 0.458, its hazard ratio is exp (0.458) =1.581. Holding other factors constant, males are (1.581-1)*100%, that is 58.1% more likely to transfer than the female.

4.4.2 Residence

The estimate for Lari was -0.25, exp (0.25) =1.284. Adjusting for other factors, teachers staying in Lari are28.4% less likely to transfer compared to those staying in Kikuyu. Limuru’s hazard ratio was exp (0.258) =1.294 this indicates that teachers staying in Limuru are 29.4% more likely to transfer than those residing in Kikuyu. Nairobi’s hazard ratio was exp (0.211) =1.235, indicating that teachers living in Nairobi are 23.5% more likely to transfer as opposed to those living in Kikuyu.

4.4.3 Education Level

Teachers with diploma qualification are exp (1.053) =2.866 times likely to transfer in comparison to the educators with Bachelor's Degree. Teachers with master's degree are exp (0.027) =1.027, 2.7% more likely to change working region compared to the ones with a bachelor's degree in education. Teachers with a doctor's degree are exp (0.045) =1.046, 4.6% less likely to transfer compared to those with a Bachelor's Degree.

4.4.4 Age

With a unit increase in age, a teacher is exp (0.032) =1.0325, 3.25% less likely to transfer.

4.4.5 Marital Status

Single teachers are (exp (0.392)-1)*100= 48% more likely to transfer compared to the married ones.

4.4.6 Teaching Subjects

Biology teachers are exp (1.458) =4.3 times more likely to transfer compared to agriculture teachers. Business studies teachers are 2.95 times more likely to transfer than agriculture teachers. Chemistry teachers are (exp (0.239)-1)100=27% more likely to transfer compared to agriculture teachers. Computer studies teachers are 3.11 times more likely to transfer compared to agriculture teachers. Christian education teachers are 5.78 times more likely to transfer than their agriculture counter parts. English teachers are (exp (0.608)-1)100= 83.5% more likely to transfer compared to agriculture teachers. Geography teachers have exp(0.945)=2.578 times more chances of transferring compared to their agriculture counterparts .Teachers teaching history are (exp(0.378)-1)100=46% less likely to transfer as opposed to those teaching agriculture. Home science teachers are (exp (0.575)-1)100=77.8% more likely to transfer compared to agriculture teachers. Kiswahili teachers are (exp (0.458)-1)100=58.1% more likely to transfer compared to agriculture teachers. Mathematics teachers have exp (0.745) =2.016 times more chances of transferring compared to their agriculture counterparts Physics teachers have exp (0.694) =2.002 times chance to transfer compared to their agriculture counterparts.

4.4.7 Present Assignment

Heads of departments are (exp (0.627)-1)100=87.2% more likely to transfer in comparison to deputy head teachers. Principals are (exp (0.061)-1)100=6.3% more likely to transfer than the deputy heads. Ordinary teachers are (exp (0.149)-1)100=16.07% less likely to transfer in relation to the deputy head teachers.

CHAPTER 5

5.0 Conclusion and Recommendations.

This study focused on the importance of gender, marital status, teaching subjects, age, education level and place of residence on teachers' decision to transfer out of a sub county. The main reasons for teacher turnover where gender, level of education, age, marital status, teaching subjects and the present assignment of a teacher. Place of residence of a teacher was not found to significantly contribute to the decision of a teacher to transfer out of a sub county. In relation to gender, male teachers were more likely to transfer in relation to the female teachers. On education level, the study concluded that teachers with diploma qualification had a higher chance of transferring while those with a doctoral degree had the least probability of transferring. For a unit increase in age, a teacher's chances of transferring decreases by 3%.Also the study concluded that single teachers have a higher chance of transferring than the married .Teacher who teaches biology, computer studies, Christian religious education, business studies and geography have a higher chance of transferring. Lastly, heads of departments had a higher chance of transferring while the deputy head teachers are the least likely to transfer. The place of residence of a teacher was not found to influence a teacher's decision to transfer. Generally the transfer rate is positively associated with the duration of stay.

In curbing teacher attrition by the government and stakeholders maneuver ways of enticing young teachers as well as providing conducive environment to encourage teacher's advancement in their studies.

REFERENCES

Ghadimi M, Mahmoodi M, Mohammad K, Zeraati H, Rasouli M and Sheikhfathollahi M. Family history of the cancer on the survival of the patients with gastrointestinal cancer in northern Iran, using frailty models. *BMC Gastroenterol*. 2011; 11:104. Article | PubMed Abstract | PubMed FullText.

John, J., Jan, o., and Emily, P., (2008). The determinants of teacher attrition in upstate New York.

Sharon, J., Betty, J. and Bernard, C. (2014). Institutional factors that influence teacher turnover in public secondary schools in Baringo district, Kenya. International journal of education and research, vol. 2 no. 4.

Hassan, A., (2013). Why secondary school teachers leave the profession in Afghanistan.

Linda, K., (1988). A study of the causes of teacher attrition in regular and special education in Wisconsin.

Herbert, K. and Ramsay, M. (2004). Teacher Turnover and Shortages of Qualified Teachers in Texas Public School Districts 2001-2004, Report to the Senate Education Committee, September 2004.

Finlayson, M. (2003).Improving the wellbeing of teachers in Scotland, Scott. Education Journal, 87(1): 18-19.

Santiago, P. (2001). Teacher shortage .Observer OECD. Retrieved May 14, 2012, from http://www.oecdobserver.org/news/printpage.php/aid/431/Teacher_shortage.html.

Xaba M. I. (2003). Factors influencing the job satisfaction of senior teachers in schools predominantly attended by black students. Med dissertation. Vanderbijlpark: Potchefstroom University.

Kadzamira, E. C. (2005) Teacher Motivation and Incentives in Malawi, Zambia: Centre for Education Research and Training, University of Malawi.1.No.11 November, 2013 pp110. Www.ijern.com.

Orodho, A.J., Waweru, P.N., Getange, K.N & Miriti, J.M.(2013).Progress towards attainment of education for All (EFA) among nomadic pastoralists : Do home-based variables make a difference in Kenya? Research on Humanities and Social Sciences.Vol3.No.21, 2013 pp54-67. www.iiste.org.

Oketch, M. &Ngware, M.W (eds.) (2012).urbanization and Education in East n Africa: African Population and Health Research Center. ISBN 978-9966-21-175-0 .

Oyaro, K. (2008). Education-Kenya: Students Pour In, Teachers Drain Away; Inter Press Service News Agency. Retrieved April18, 2011, from http://ipsnews.net/africa/nota.asp?idnews=42667.

Ondara, O. K. (2004). An investigation into job satisfaction among secondary school teachers: A case study of Borabu Division in Nyamira District, Kenya. Unpublished Master of Education Thesis, Kenyatta University.

Oplatka, I. (2007). The context and profile of teachers in developing countries in the last decade: A revealing discussion for further investigation, International Journal of Educational Management, 21 (6), 476-490.

APPENDIX

R codes

```
setwd("C:/Users/ivan/Desktop/Data g Final")#data generation
set.seed(123)
subcounty<-rep(1:10,300)
set.seed(89)
gend<-round(runif(300, 1,2),0)
gend[gend==1]<-c("male")
gend[gend==2]<-c("female")
gender<-gend
set.seed(904)
plac<-round(runif(300, 1,4),0)
plac[plac==1]<-c("Lari")
plac[plac==2]<-c("Nairobi")
plac[plac==3]<-c("Kikuyu")
plac[plac==4]<-c("Limuru")
residence<-plac
set.seed(670)
edc<-round(runif(300, 1,4),0)
edc[edc==1]<-c("Diploma")
edc[edc==2]<-c("Bed")
edc[edc==3]<-c("MSc")
edc[edc==4]<-c("PhD")
education_level<-edc
education_level
set.seed(594)
ts<-round(runif(300, 1,13),0)
ts[ts==1]<-c("Home_science")
ts[ts==2]<-c("Mathematics")
```

```
ts[ts==3]<-c("English")
ts[ts==4]<-c("Kiswahili")
ts[ts==5]<-c("CRE")
ts[ts==6]<-c("Geography")
ts[ts==7]<-c("Agriculture")
ts[ts==8]<-c("Business_studies")
ts[ts==9]<-c("History")
ts[ts==10]<-c("Physics")
ts[ts==11]<-c("Biology")
ts[ts==12]<-c("Chemistry")
ts[ts==13]<-c("Computer_Studies")
teaching_subjects<-ts
set.seed(346)
pa<-round(runif(300, 1,4),0)
pa[pa==1]<-c("HoD")
pa[pa==2]<-c("Principal")
pa[pa==3]<-c("Deputy_head")
pa[pa==4]<-c("Teacher")
Present_Assignment<-pa
set.seed(7893)
Age<-round(runif(300,24,60),0)
set.seed(76609)
mar<-round(runif(300,1,2),0)
mar[mar==1]<-c("Married")
mar[mar==2]<-c("Single")
Marital_status<-mar

set.seed(9547)
schcategory<-round(runif(300,1,4),0)
```

```
schcategory[schcategory==1]<-c("National")

schcategory[schcategory==2]<-c("County")

schcategory[schcategory==3]<-c("Extra_county")

schcategory[schcategory==4]<-c("CDF")

school_category<-schcategory


set.seed(76325)

duration_stay<-round(runif(300,1,20),0)

attri_data<-
as.data.frame(cbind(subcounty,gender,residence,education_level,teaching_subjects,Present_
Assignment,Age,Marital_status,school_category,duration_stay))

View(attri_data)# to see the data

colnames(attri_data)<-
c("Subcounty","Gender","Residence","Educationlevel","Teachingsubjects","Presentassignme
nt","Age","Maritalstatus","Schoolcategory","Durationstay")

write.table(attri_data,file = "attri_data_2.csv",sep=",")#writing the data

datacc<-read.csv("attri_data_2.csv",sep=",",header=TRUE) # to read the data before
attaching

attach(datacc)

names(datacc)

datacc$sampledreasons<-
ifelse(Age<=50,sample(c("resignation","death","sacking","transfer"),300,replace=TRUE),sa
mple(c("resignation","death","sacking","retirement","transfer"),300,replace=TRUE))

View(datacc)

write.table(datacc,file = "datacc1.csv",sep=",")#writing the data

datacc1<-read.csv("datacc1.csv",sep=",",header=TRUE) # to read the data before attaching

attach(datacc1)

sampledreasons1<-ifelse(sampledreasons=="transfer",1,0)

sampledreasons1

library(parfm)

Sv<-Surv(Durationstay,sampledreasons1)

Sv
```

```
plot(Sv,ylab = "H(t)")#plotting hazard fn
attrition.mod1 <- parfm(Surv(Durationstay, sampledreasons1) ~ 1,
                    cluster ="Subcounty",
                    dist = "weibull",
                    frailty = "gamma",
                    data = datacc1,
                    iniFpar=0.001)
attrition.mod2 <- parfm(Surv(Durationstay, sampledreasons1) ~ (Gender),
                    cluster ="Subcounty",
                    dist = "weibull",
                    frailty = "gamma",
                    data = datacc1,
                    iniFpar=0.001)
attrition.mod3 <- parfm(Surv(Durationstay, sampledreasons1) ~ ( Residence),
                    cluster ="Subcounty",
                    dist = "weibull",
                    frailty = "gamma",
                    data = datacc1,
                    iniFpar=0.001)
attrition.mod4 <- parfm(Surv(Durationstay, sampledreasons1) ~ (Educationlevel),
                    cluster ="Subcounty",
                    dist = "weibull",
                    frailty = "gamma",
                    data = datacc1,
                    iniFpar=0.001)
attrition.mod5 <- parfm(Surv(Durationstay, sampledreasons1) ~ (Age),
                    cluster ="Subcounty",
                    dist = "weibull",
                    frailty = "gamma",
```

```
                     data = datacc1,
                     iniFpar=0.001)
attrition.mod6 <- parfm(Surv(Durationstay, sampledreasons1) ~ (Maritalstatus),
                     cluster ="Subcounty",
                     dist = "weibull",
                     frailty = "gamma",
                     data = datacc1,
                     iniFpar=0.001)
attrition.mod7 <- parfm(Surv(Durationstay, sampledreasons1) ~ ( Teachingsubjects),
                     cluster ="Subcounty",
                     dist = "weibull",
                     frailty = "gamma",
                     data = datacc1,
                     iniFpar=0.001)
attrition.mod8 <- parfm(Surv(Durationstay, sampledreasons1) ~ (Schoolcategory),
                     cluster ="Subcounty",
                     dist = "weibull",
                     frailty = "gamma",
                     data = datacc1,
                     iniFpar=0.001)
attrition.mod9 <- parfm(Surv(Durationstay, sampledreasons1) ~ (Presentassignment),
                     cluster ="Subcounty",
                     dist = "weibull",
                     frailty = "gamma",
                     data = datacc1,
                     iniFpar=0.001)
attrition.mod10 <- parfm(Surv(Durationstay, sampledreasons1) ~ (subcounty),
                      cluster ="Subcounty",
                      dist = "weibull",
```

```
                    frailty = "gamma",
                    data = datacc1,
                    iniFpar=0.001)
attrition.mod11 <- parfm(Surv(Durationstay, sampledreasons1) ~ (Gender + Residence),
                    cluster ="Subcounty",
                    dist = "weibull",
                    frailty = "gamma",
                    data = datacc1,
                    iniFpar=0.001)
attrition.mod12 <- parfm(Surv(Durationstay, sampledreasons1) ~ (Gender+Educationlevel),
                    cluster ="Subcounty",
                    dist = "weibull",
                    frailty = "gamma",
                    data = datacc1,
                    iniFpar=0.001)
attrition.mod13 <- parfm(Surv(Durationstay, sampledreasons1) ~ (Gender+ Age),
                    cluster ="Subcounty",
                    dist = "weibull",
                    frailty = "gamma",
                    data = datacc1,
                    iniFpar=0.001)
attrition.mod14 <- parfm(Surv(Durationstay, sampledreasons1) ~ (Gender+Maritalstatus),
                    cluster ="Subcounty",
                    dist = "weibull",
                    frailty = "gamma",
                    data = datacc1,
                    iniFpar=0.001)
attrition.mod15 <- parfm(Surv(Durationstay, sampledreasons1) ~
(Gender+Teachingsubjects),
```

```
                cluster ="Subcounty",
                dist = "weibull",
                frailty = "gamma",
                data = datacc1,
                iniFpar=0.001)
attrition.mod16 <- parfm(Surv(Durationstay, sampledreasons1) ~ (Gender+Schoolcategory),
                cluster ="Subcounty",
                dist = "weibull",
                frailty = "gamma",
                data = datacc1,
                iniFpar=0.001)
attrition.mod17 <- parfm(Surv(Durationstay, sampledreasons1) ~
(Gender+Presentassignment),
                cluster ="Subcounty",
                dist = "weibull",
                frailty = "gamma",
                data = datacc1,
                iniFpar=0.001)
attrition.mod18 <- parfm(Surv(Durationstay, sampledreasons1) ~
(Gender+Presentassignment),
                cluster ="Subcounty",
                dist = "weibull",
                frailty = "gamma",
                data = datacc1,
                iniFpar=0.001)
attrition.mod19 <- parfm(Surv(Durationstay, sampledreasons1) ~ (Gender+subcounty),
                cluster ="Subcounty",
                dist = "weibull",
                frailty = "gamma",
                data = datacc1,
```

```
                    iniFpar=0.001)
attrition.mod20 <- parfm(Surv(Durationstay, sampledreasons1) ~ (Residence +
Educationlevel),
                    cluster ="Subcounty",
                    dist = "weibull",
                    frailty = "gamma",
                    data = datacc1,
                    iniFpar=0.001)
attrition.mod21 <- parfm(Surv(Durationstay, sampledreasons1) ~ (Residence +Age),
                    cluster ="Subcounty",
                    dist = "weibull",
                    frailty = "gamma",
                    data = datacc1,
                    iniFpar=0.001)
attrition.mod22 <- parfm(Surv(Durationstay, sampledreasons1) ~ (Residence +
Maritalstatus),
                    cluster ="Subcounty",
                    dist = "weibull",
                    frailty = "gamma",
                    data = datacc1,
                    iniFpar=0.001)
attrition.mod23 <- parfm(Surv(Durationstay, sampledreasons1) ~ (Residence +
Teachingsubjects),
                    cluster ="Subcounty",
                    dist = "weibull",
                    frailty = "gamma",
                    data = datacc1,
                    iniFpar=0.001)
attrition.mod24 <- parfm(Surv(Durationstay, sampledreasons1) ~ (Residence
+Schoolcategory),
```

```
                    cluster ="Subcounty",
                    dist = "weibull",
                    frailty = "gamma",
                    data = datacc1,
                    iniFpar=0.001)
attrition.mod25 <- parfm(Surv(Durationstay, sampledreasons1) ~ (Residence + 
Presentassignment),
                    cluster ="Subcounty",
                    dist = "weibull",
                    frailty = "gamma",
                    data = datacc1,
                    iniFpar=0.001)
attrition.mod26 <- parfm(Surv(Durationstay, sampledreasons1) ~ (Educationlevel + Age),
                    cluster ="Subcounty",
                    dist = "weibull",
                    frailty = "gamma",
                    data = datacc1,
                    iniFpar=0.001)
attrition.mod27 <- parfm(Surv(Durationstay, sampledreasons1) ~ 
(Educationlevel+Maritalstatus),
                    cluster ="Subcounty",
                    dist = "weibull",
                    frailty = "gamma",
                    data = datacc1,
                    iniFpar=0.001)
attrition.mod28 <- parfm(Surv(Durationstay, sampledreasons1) ~ (
Educationlevel+Teachingsubjects),
                    cluster ="Subcounty",
                    dist = "weibull",
                    frailty = "gamma",
```

```
              data = datacc1,
              iniFpar=0.001)
attrition.mod29 <- parfm(Surv(Durationstay, sampledreasons1) ~ (Educationlevel
+Schoolcategory),
              cluster ="Subcounty",
              dist = "weibull",
              frailty = "gamma",
              data = datacc1,
              iniFpar=0.001)
attrition.mod30 <- parfm(Surv(Durationstay, sampledreasons1) ~ ( Educationlevel
+Presentassignment),
              cluster ="Subcounty",
              dist = "weibull",
              frailty = "gamma",
              data = datacc1,
              iniFpar=0.001)
attrition.mod31 <- parfm(Surv(Durationstay, sampledreasons1) ~ (Age + Maritalstatus),
              cluster ="Subcounty",
              dist = "weibull",
              frailty = "gamma",
              data = datacc1,
              iniFpar=0.001)
attrition.mod32 <- parfm(Surv(Durationstay, sampledreasons1) ~ (Age +Teachingsubjects),
              cluster ="Subcounty",
              dist = "weibull",
              frailty = "gamma",
              data = datacc1,
              iniFpar=0.001)
attrition.mod33 <- parfm(Surv(Durationstay, sampledreasons1) ~ (Age +Schoolcategory),
              cluster ="Subcounty",
```

```
                    dist = "weibull",
                    frailty = "gamma",
                    data = datacc1,
                    iniFpar=0.001)
attrition.mod34 <- parfm(Surv(Durationstay, sampledreasons1) ~ (Age +Presentassignment),
                    cluster ="Subcounty",
                    dist = "weibull",
                    frailty = "gamma",
                    data = datacc1,
                    iniFpar=0.001)
attrition.mod35 <- parfm(Surv(Durationstay, sampledreasons1) ~
(Maritalstatus+Teachingsubjects),
                    cluster ="Subcounty",
                    dist = "weibull",
                    frailty = "gamma",
                    data = datacc1,
                    iniFpar=0.001)
attrition.mod36 <- parfm(Surv(Durationstay, sampledreasons1) ~
(Maritalstatus+Schoolcategory),
                    cluster ="Subcounty",
                    dist = "weibull",
                    frailty = "gamma",
                    data = datacc1,
                    iniFpar=0.001)
attrition.mod37 <- parfm(Surv(Durationstay, sampledreasons1) ~
(Maritalstatus+Presentassignment),
                    cluster ="Subcounty",
                    dist = "weibull",
                    frailty = "gamma",
                    data = datacc1,
```

```
                iniFpar=0.001)
attrition.mod38 <- parfm(Surv(Durationstay, sampledreasons1) ~
(Teachingsubjects+Schoolcategory),
                cluster ="Subcounty",
                dist = "weibull",
                frailty = "gamma",
                data = datacc1,
                iniFpar=0.001)
attrition.mod39 <- parfm(Surv(Durationstay, sampledreasons1) ~
(Teachingsubjects+Presentassignment),
                cluster ="Subcounty",
                dist = "weibull",
                frailty = "gamma",
                data = datacc1,
                iniFpar=0.001)
attrition.mod40 <- parfm(Surv(Durationstay, sampledreasons1) ~
(Teachingsubjects+Schoolcategory),
                cluster ="Subcounty",
                dist = "weibull",
                frailty = "gamma",
                data = datacc1,
                iniFpar=0.001)
attrition.mod41 <- parfm(Surv(Durationstay, sampledreasons1) ~
(Teachingsubjects+Presentassignment),
                cluster ="Subcounty",
                dist = "weibull",
                frailty = "gamma",
                data = datacc1,
                iniFpar=0.001)
attrition.mod42 <- parfm(Surv(Durationstay, sampledreasons1) ~ (Gender + Residence +
Educationlevel),
```

```
                    cluster ="Subcounty",
                    dist = "weibull",
                    frailty = "gamma",
                    data = datacc1,
                    iniFpar=0.001)
attrition.mod43 <- parfm(Surv(Durationstay, sampledreasons1) ~ (Gender + Residence +
Age),
                    cluster ="Subcounty",
                    dist = "weibull",
                    frailty = "gamma",
                    data = datacc1,
                    iniFpar=0.001)
attrition.mod44 <- parfm(Surv(Durationstay, sampledreasons1) ~ (Gender + Residence
+Maritalstatus),
                    cluster ="Subcounty",
                    dist = "weibull",
                    frailty = "gamma",
                    data = datacc1,
                    iniFpar=0.001)
attrition.mod45 <- parfm(Surv(Durationstay, sampledreasons1) ~ (Gender + Residence
+Teachingsubjects),
                    cluster ="Subcounty",
                    dist = "weibull",
                    frailty = "gamma",
                    data = datacc1,
                    iniFpar=0.001)
attrition.mod46 <- parfm(Surv(Durationstay, sampledreasons1) ~ (Gender + Residence +
Schoolcategory),
                    cluster ="Subcounty",
                    dist = "weibull",
```

```
                    frailty = "gamma",
                    data = datacc1,
                    iniFpar=0.001)
attrition.mod47 <- parfm(Surv(Durationstay, sampledreasons1) ~ (Gender + Residence +
Presentassignment),
                    cluster ="Subcounty",
                    dist = "weibull",
                    frailty = "gamma",
                    data = datacc1,
                    iniFpar=0.001)
attrition.mod48 <- parfm(Surv(Durationstay, sampledreasons1) ~ (Residence +
Educationlevel + Age),
                    cluster ="Subcounty",
                    dist = "weibull",
                    frailty = "gamma",
                    data = datacc1,
                    iniFpar=0.001)
attrition.mod49 <- parfm(Surv(Durationstay, sampledreasons1) ~ (Residence +
Educationlevel +Maritalstatus),
                    cluster ="Subcounty",
                    dist = "weibull",
                    frailty = "gamma",
                    data = datacc1,
                    iniFpar=0.001)
attrition.mod50 <- parfm(Surv(Durationstay, sampledreasons1) ~ (Residence +
Educationlevel +Teachingsubjects),
                    cluster ="Subcounty",
                    dist = "weibull",
                    frailty = "gamma",
                    data = datacc1,
```

```
                 iniFpar=0.001)
attrition.mod51 <- parfm(Surv(Durationstay, sampledreasons1) ~ (Residence +
Educationlevel + Schoolcategory),
                 cluster ="Subcounty",
                 dist = "weibull",
                 frailty = "gamma",
                 data = datacc1,
                 iniFpar=0.001)
attrition.mod52 <- parfm(Surv(Durationstay, sampledreasons1) ~ (Residence +
Educationlevel +Presentassignment),
                 cluster ="Subcounty",
                 dist = "weibull",
                 frailty = "gamma",
                 data = datacc1,
                 iniFpar=0.001)
attrition.mod53 <- parfm(Surv(Durationstay, sampledreasons1) ~ (Educationlevel + Age +
Maritalstatus),
                 cluster ="Subcounty",
                 dist = "weibull",
                 frailty = "gamma",
                 data = datacc1,
                 iniFpar=0.001)
attrition.mod54 <- parfm(Surv(Durationstay, sampledreasons1) ~ (Educationlevel + Age +
Teachingsubjects),
                 cluster ="Subcounty",
                 dist = "weibull",
                 frailty = "gamma",
                 data = datacc1,
                 iniFpar=0.001)
attrition.mod55 <- parfm(Surv(Durationstay, sampledreasons1) ~ (Educationlevel + Age +
Schoolcategory),
```

```
cluster ="Subcounty",
dist = "weibull",
frailty = "gamma",
data = datacc1,
iniFpar=0.001)
attrition.mod56 <- parfm(Surv(Durationstay, sampledreasons1) ~ (Educationlevel + Age +Presentassignment),
cluster ="Subcounty",
dist = "weibull",
frailty = "gamma",
data = datacc1,
iniFpar=0.001)
attrition.mod57 <- parfm(Surv(Durationstay, sampledreasons1) ~ (Age + Maritalstatus+Teachingsubjects),
cluster ="Subcounty",
dist = "weibull",
frailty = "gamma",
data = datacc1,
iniFpar=0.001)
attrition.mod58 <- parfm(Surv(Durationstay, sampledreasons1) ~ (Age + Maritalstatus+Schoolcategory),
cluster ="Subcounty",
dist = "weibull",
frailty = "gamma",
data = datacc1,
iniFpar=0.001)
attrition.mod59 <- parfm(Surv(Durationstay, sampledreasons1) ~ (Age + Maritalstatus+Presentassignment),
cluster ="Subcounty",
dist = "weibull",
```

```
                    frailty = "gamma",
                    data = datacc1,
                    iniFpar=0.001)
attrition.mod60 <- parfm(Surv(Durationstay, sampledreasons1) ~
(Maritalstatus+Teachingsubjects+Schoolcategory),
                    cluster ="Subcounty",
                    dist = "weibull",
                    frailty = "gamma",
                    data = datacc1,
                    iniFpar=0.001)
attrition.mod61 <- parfm(Surv(Durationstay, sampledreasons1) ~
(Maritalstatus+Teachingsubjects+Presentassignment),
                    cluster ="Subcounty",
                    dist = "weibull",
                    frailty = "gamma",
                    data = datacc1,
                    iniFpar=0.001)
attrition.mod62 <- parfm(Surv(Durationstay, sampledreasons1) ~ (Gender + Residence +
Educationlevel + Age),
                    cluster ="Subcounty",
                    dist = "weibull",
                    frailty = "gamma",
                    data = datacc1,
                    iniFpar=0.001)
attrition.mod63 <- parfm(Surv(Durationstay, sampledreasons1) ~ (Gender + Residence +
Educationlevel + Maritalstatus),
                    cluster ="Subcounty",
                    dist = "weibull",
                    frailty = "gamma",
                    data = datacc1,
```

```
              iniFpar=0.001)
attrition.mod64 <- parfm(Surv(Durationstay, sampledreasons1) ~ (Gender + Residence + Educationlevel + Teachingsubjects),
              cluster ="Subcounty",
              dist = "weibull",
              frailty = "gamma",
              data = datacc1,
              iniFpar=0.001)
attrition.mod65 <- parfm(Surv(Durationstay, sampledreasons1) ~ (Gender + Residence + Educationlevel +Schoolcategory),
              cluster ="Subcounty",
              dist = "weibull",
              frailty = "gamma",
              data = datacc1,
              iniFpar=0.001)
attrition.mod66 <- parfm(Surv(Durationstay, sampledreasons1) ~ (Gender + Residence + Educationlevel + Presentassignment),
              cluster ="Subcounty",
              dist = "weibull",
              frailty = "gamma",
              data = datacc1,
              iniFpar=0.001)
attrition.mod67 <- parfm(Surv(Durationstay, sampledreasons1) ~ (Residence + Educationlevel + Age + Maritalstatus),
              cluster ="Subcounty",
              dist = "weibull",
              frailty = "gamma",
              data = datacc1,
              iniFpar=0.001)
attrition.mod68 <- parfm(Surv(Durationstay, sampledreasons1) ~ (Residence + Educationlevel + Age + Teachingsubjects),
```

```
                    cluster ="Subcounty",
                    dist = "weibull",
                    frailty = "gamma",
                    data = datacc1,
                    iniFpar=0.001)
attrition.mod69 <- parfm(Surv(Durationstay, sampledreasons1) ~ (Residence +
Educationlevel + Age + Schoolcategory),
                    cluster ="Subcounty",
                    dist = "weibull",
                    frailty = "gamma",
                    data = datacc1,
                    iniFpar=0.001)


attrition.mod70 <- parfm(Surv(Durationstay, sampledreasons1) ~ (Residence +
Educationlevel + Age + Presentassignment),
                    cluster ="Subcounty",
                    dist = "weibull",
                    frailty = "gamma",
                    data = datacc1,
                    iniFpar=0.001)


attrition.mod71 <- parfm(Surv(Durationstay, sampledreasons1) ~ (Educationlevel + Age +
Maritalstatus+Teachingsubjects),
                    cluster ="Subcounty",
                    dist = "weibull",
                    frailty = "gamma",
                    data = datacc1,
                    iniFpar=0.001)
attrition.mod72 <- parfm(Surv(Durationstay, sampledreasons1) ~ (Educationlevel + Age +
Maritalstatus+Schoolcategory),
```

```
              cluster ="Subcounty",
              dist = "weibull",
              frailty = "gamma",
              data = datacc1,
              iniFpar=0.001)
attrition.mod73 <- parfm(Surv(Durationstay, sampledreasons1) ~ (Educationlevel + Age +
Maritalstatus+Presentassignment),
              cluster ="Subcounty",
              dist = "weibull",
              frailty = "gamma",
              data = datacc1,
              iniFpar=0.001)
attrition.mod74 <- parfm(Surv(Durationstay, sampledreasons1) ~ (Age +
Maritalstatus+Teachingsubjects+Schoolcategory),
              cluster ="Subcounty",
              dist = "weibull",
              frailty = "gamma",
              data = datacc1,
              iniFpar=0.001)
attrition.mod75 <- parfm(Surv(Durationstay, sampledreasons1) ~ (Age +
Maritalstatus+Teachingsubjects+Presentassignment),
              cluster ="Subcounty",
              dist = "weibull",
              frailty = "gamma",
              data = datacc1,
              iniFpar=0.001)



attrition.mod76 <- parfm(Surv(Durationstay, sampledreasons1) ~ (Gender + Residence +
Educationlevel + Age + Maritalstatus),
```

```
                    cluster ="Subcounty",
                    dist = "weibull",
                    frailty = "gamma",
                    data = datacc1,
                    iniFpar=0.001)
attrition.mod77 <- parfm(Surv(Durationstay, sampledreasons1) ~ (Gender + Residence +
Educationlevel + Age + Teachingsubjects),
                    cluster ="Subcounty",
                    dist = "weibull",
                    frailty = "gamma",
                    data = datacc1,
                    iniFpar=0.001)
attrition.mod78 <- parfm(Surv(Durationstay, sampledreasons1) ~ (Gender + Residence +
Educationlevel + Age + Schoolcategory),
                    cluster ="Subcounty",
                    dist = "weibull",
                    frailty = "gamma",
                    data = datacc1,
                    iniFpar=0.001)
attrition.mod79 <- parfm(Surv(Durationstay, sampledreasons1) ~ (Gender + Residence +
Educationlevel + Age + Presentassignment),
                    cluster ="Subcounty",
                    dist = "weibull",
                    frailty = "gamma",
                    data = datacc1,
                    iniFpar=0.001)
attrition.mod80 <- parfm(Surv(Durationstay, sampledreasons1) ~ (Residence +
Educationlevel + Age + Maritalstatus+Teachingsubjects),
                    cluster ="Subcounty",
                    dist = "weibull",
```

```
                    frailty = "gamma",

                    data = datacc1,

                    iniFpar=0.001)

attrition.mod81 <- parfm(Surv(Durationstay, sampledreasons1) ~ (Residence +
Educationlevel + Age + Maritalstatus+Schoolcategory),

                    cluster ="Subcounty",

                    dist = "weibull",

                    frailty = "gamma",

                    data = datacc1,

                    iniFpar=0.001)

attrition.mod82 <- parfm(Surv(Durationstay, sampledreasons1) ~ (Residence +
Educationlevel + Age + Maritalstatus+Presentassignment),

                    cluster ="Subcounty",

                    dist = "weibull",

                    frailty = "gamma",

                    data = datacc1,

                    iniFpar=0.001)

attrition.mod83 <- parfm(Surv(Durationstay, sampledreasons1) ~ (Educationlevel + Age +
Maritalstatus+Teachingsubjects+Schoolcategory),

                    cluster ="Subcounty",

                    dist = "weibull",

                    frailty = "gamma",

                    data = datacc1,

                    iniFpar=0.001)

attrition.mod84 <- parfm(Surv(Durationstay, sampledreasons1) ~ (Educationlevel + Age +
Maritalstatus+Teachingsubjects+Presentassignment),

                    cluster ="Subcounty",

                    dist = "weibull",

                    frailty = "gamma",

                    data = datacc1,
```

```
                iniFpar=0.001)
attrition.mod85 <- parfm(Surv(Durationstay, sampledreasons1) ~ (Gender + Residence +
Educationlevel + Age + Maritalstatus+Teachingsubjects),
                cluster ="Subcounty",
                dist = "weibull",
                frailty = "gamma",
                data = datacc1,
                iniFpar=0.001)
attrition.mod86 <- parfm(Surv(Durationstay, sampledreasons1) ~ (Gender + Residence +
Educationlevel + Age + Maritalstatus+Schoolcategory),
                cluster ="Subcounty",
                dist = "weibull",
                frailty = "gamma",
                data = datacc1,
                iniFpar=0.001)
attrition.mod87 <- parfm(Surv(Durationstay, sampledreasons1) ~ (Gender + Residence +
Educationlevel + Age + Maritalstatus+Presentassignment),
                cluster ="Subcounty",
                dist = "weibull",
                frailty = "gamma",
                data = datacc1,
                iniFpar=0.001)
attrition.mod88 <- parfm(Surv(Durationstay, sampledreasons1) ~ (Residence +
Educationlevel + Age + Maritalstatus+Teachingsubjects+Schoolcategory),
                cluster ="Subcounty",
                dist = "weibull",
                frailty = "gamma",
                data = datacc1,
                iniFpar=0.001)
attrition.mod89 <- parfm(Surv(Durationstay, sampledreasons1) ~ (Residence +
Educationlevel + Age + Maritalstatus+Teachingsubjects+Presentassignment),
```

```
                cluster ="Subcounty",
                dist = "weibull",
                frailty = "gamma",
                data = datacc1,
                iniFpar=0.001)
attrition.mod90 <- parfm(Surv(Durationstay, sampledreasons1) ~ (Gender + Residence + Educationlevel + Age + Maritalstatus+Teachingsubjects+Schoolcategory),
                cluster ="Subcounty",
                dist = "weibull",
                frailty = "gamma",
                data = datacc1,
                iniFpar=0.001)
attrition.mod91 <- parfm(Surv(Durationstay, sampledreasons1) ~ (Gender + Residence + Educationlevel + Age + Maritalstatus+Teachingsubjects+Presentassignment),
                cluster ="Subcounty",
                dist = "weibull",
                frailty = "gamma",
                data = datacc1,
                iniFpar=0.001)
attrition.mod92 <- parfm(Surv(Durationstay, sampledreasons1) ~ (Residence + Educationlevel + Age + Maritalstatus+Teachingsubjects+Schoolcategory+Presentassignment),
                cluster ="Subcounty",
                dist = "weibull",
                frailty = "gamma",
                data = datacc1,
                iniFpar=0.001)
attrition.mod93 <- parfm(Surv(Durationstay, sampledreasons1) ~ (Gender + Residence + Educationlevel + Age + Maritalstatus+Teachingsubjects+Schoolcategory+Presentassignment),
                cluster ="Subcounty",
```

```
                 dist = "weibull",
                 frailty = "gamma",
                 data = datacc1,
                 iniFpar=0.001)
attrition.mod93
attrition.mod91
datag<-read.csv(file.choose(),header = TRUE)
attach(datag)
tapply(Age,residence,mean)# analysing age by residence
boxplot(Age,duration_stay)#checking if data is normally dist.
xtabs(~Gender)#cross tabs
xtabs(~Gender+Teachingsubjects)
xtabs(~Gender+sampledreasons)
xtabs(~Residence)
xtabs(~Residence+sampledreasons)
xtabs(~Educationlevel)
xtabs(~Educationlevel+sampledreasons)
xtabs(~ Maritalstatus)
xtabs(~ Maritalstatus+sampledreasons)
xtabs(~ Presentassignment)
xtabs(~ Presentassignment+sampledreasons)
xtabs(~Teachingsubjects)
xtabs(~Teachingsubjects+sampledreasons)
xtabs(~Age)
xtabs(~Age+sampledreasons)
xtabs(~Schoolcategory)
xtabs(~Schoolcategory+sampledreasons)
```

Table 3: Cross Tabulations by Gender

Female	**Male**
1450	1550
48	52

Table 4: Cross Tabulations by gender and teaching subjects

Gender	Agriculture	Biology	Business studies	Chemistry	Computer Studies	CRE	English	Geography	History
female	48	75	42	64	36	42	37	46	45
male	52	25	58	36	64	58	63	54	55

Gender	Home_science	Kiswahili	Mathematics	Physics
female	55	42	43	56
male	45	58	57	44

Table 5: Cross Tabulations by gender and sampled reasons

Gender	Death	Resignation	Retirement	Sacking	Transfer
female	40	63	50	49	43
male	60	37	50	51	57

Table 6: Cross Tabulations by Residence and sampled reasons

Residence	death	resignation	retirement	sacking	transfer
Kikuyu	36	34	44	37	29
Lari	18	15	13	19	10
Limuru	14	14	13	16	22
Nairobi	33	37	31	29	39

Table 7: Cross Tabulations by Qualification

Bed	Diploma	MSc	PhD
1000	520	1030	450
33	17	34	15

Table 8: Cross Tabulations by Qualification and sampled reasons

Education Level	Death	Resignation	Retirement	Sacking	Transfer
Bed	33	35	56	31	29
Diploma	16	12	13	10	32
MSc	35	35	19	43	28
PhD	16	17	13	16	12

Table 9: Cross Tabulations by marital status

Married	Single
1480	1520
49	51

Table 10: Cross Tabulations by marital status and sampled reasons

Marital Status	Death	Resignation	Retirement	Sacking	Transfer
Married	43	55	38	63	41
Single	58	45	63	37	59

Table 11: Cross Tabulations by current assignment

Deputy_head	HoD	Principal	Teacher
980	470	980	570
33	16	33	19

Table 12: Cross Tabulations by present assignment and sampled reasons

Present Assignment	Death	Resignation	Retirement	Sacking	Transfer
Deputy_head	38	31	31	33	29
HoD	16	18	13	13	16
Principal	29	34	38	34	33
Teacher	18	17	19	20	22

Table 13: Cross Tabulations by teaching subject

Agric	Biology	Business studies	Chem	Computer Studies	CRE	English	Geog	History	Home science	Kiswa	Maths	Physics
230	160	190	280	140	190	270	280	220	110	310	280	340
8	5	6	9	5	6	9	9	7	4	10	9	11

Table 14: Cross Tabulations by teaching subject and sampled reasons

Teaching Subjects	**Death**	**Resignation**	**Retirement**	**Sacking**	**Transfer**
Agriculture	13	8	13	3	6
Biology	3	5	13	7	6
Business_studies	5	9	6	6	6
Chemistry	15	6	0	11	6
Computer Studies	3	3	0	6	9
CRE	5	5	0	3	14
English	8	8	31	6	10
Geography	8	11	13	7	12
History	9	11	0	9	3
Home_science	6	2	6	4	1
Kiswahili	8	12	6	16	7
Mathematics	11	8	6	7	12
Physics	9	14	6	16	9

Table 15: Cross Tabulations by age and sampled reasons

Age	**Death**	**Resignation**	**Retirement**	**Sacking**	**Transfer**
2	3	0	0	3	1
2	1	3	0	6	4
2	1	5	0	4	4
2	1	3	0	1	3
2	1	0	0	0	4

2	3	2	0	4	1
2	4	3	0	4	1
2	0	3	0	0	0
2	1	2	0	1	3
2	1	3	0	4	1
2	4	5	0	3	3
2	0	3	0	0	7
2	6	6	0	3	4
2	6	2	0	3	4
2	3	5	0	3	3
3	3	2	0	6	1
3	3	2	0	4	6
3	5	2	0	3	3
3	1	2	0	3	4
3	6	5	0	3	3
3	1	5	0	3	0
3	3	3	0	3	4
3	0	6	0	3	3
3	3	2	0	4	1
3	3	0	0	3	3
3	3	3	0	1	1
3	4	3	0	3	0
3	3	2	19	1	1
3	6	5	13	3	4
3	4	3	6	4	3
3	1	3	6	0	1
4	5	2	19	3	3
4	5	3	0	4	4
4	4	2	6	3	3
4	3	3	13	0	1
4	1	3	19	1	1
4	1	0	0	1	0

Table 16: Cross Tabulations by school category

CDF	County	Extra county	National
640	900	1110	350
21	30	37	12

Table 17: Cross Tabulations by school category and sampled reasons

School Category	Death	Resignation	Retirement	Sacking	Transfer
CDF	16	28	25	19	23
County	36	28	6	27	33
Extra county	36	32	63	40	33
National	11	12	6	14	10

Table 18: Model selection based on AIC

model	df	AIC	model	df	AIC	model	df	AIC
attrition.mod1	3	5994.903	attrition.mod32	16	5918.25	attrition.mod63	11	5842.028
attrition.mod2	4	5976.077	attrition.mod33	7	5953.169	attrition.mod64	22	5768.677
attrition.mod3	6	5969.215	attrition.mod34	7	5947.342	attrition.mod65	13	5847.592
attrition.mod4	6	5903.404	attrition.mod35	16	5887.618	attrition.mod66	13	5864.195
attrition.mod5	4	5953.137	attrition.mod36	7	5983.494	attrition.mod67	11	5849.367
attrition.mod6	4	5981.57	attrition.mod37	7	5979.716	attrition.mod68	22	5752.212
attrition.mod7	15	5913.026	attrition.mod38	18	5914.985	attrition.mod69	13	5850.616
attrition.mod8	6	5997.796	attrition.mod39	18	5893.913	attrition.mod70	13	5846.852
attrition.mod9	6	5991.794	attrition.mod40	18	5914.985	attrition.mod71	20	5744.135
attrition.mod10	4	5994.369	attrition.mod41	18	5893.913	attrition.mod72	11	5859.688
attrition.mod11	7	5946.311	attrition.mod42	10	5855.078	attrition.mod73	11	5853.856
attrition.mod12	7	5870.851	attrition.mod43	8	5905.534	attrition.mod74	20	5844.667
attrition.mod13	5	5931.23	attrition.mod44	8	5926.637	attrition.mod75	20	5822.357
attrition.mod14	5	5963.177	attrition.mod45	19	5888.423	attrition.mod76	12	5807.433
attrition.mod15	16	5904.998	attrition.mod46	10	5950.144	attrition.mod77	23	5723.924
attrition.mod16	7	5979.884	attrition.mod47	10	5944.793	attrition.mod78	14	5815.15
attrition.mod17	7	5975.232	attrition.mod48	10	5857.889	attrition.mod79	14	5813.939
attrition.mod18	7	5975.232	attrition.mod49	10	5879.257	attrition.mod80	23	5736.493
attrition.mod19	5	5974.985	attrition.mod50	21	5791.312	attrition.mod81	14	5842.307
attrition.mod20	9	5889.458	attrition.mod51	12	5879.244	attrition.mod82	14	5839.108
attrition.mod21	7	5932.022	attrition.mod52	12	5883.249	attrition.mod83	23	5743.245
attrition.mod22	7	5950.688	attrition.mod53	8	5862.143	attrition.mod84	23	5711.173
attrition.mod23	18	5897.854	attrition.mod54	19	5758.401	attrition.mod85	24	5703.996
attrition.mod24	9	5970.758	attrition.mod55	10	5865.439	attrition.mod86	15	5803.87
attrition.mod25	9	5963.973	attrition.mod56	10	5858.768	attrition.mod87	15	5803.539
attrition.mod26	7	5867.847	attrition.mod57	17	5847.488	attrition.mod88	26	5732.711
attrition.mod27	7	5897.343	attrition.mod58	8	5937.543	attrition.mod89	26	5705.477
attrition.mod28	18	5797.106	attrition.mod59	8	5934.113	attrition.mod90	27	5705.468
attrition.mod29	9	5899.828	attrition.mod60	19	5888.423	attrition.mod91	27	5682.665
attrition.mod30	9	5898.315	attrition.mod61	19	5871.484	attrition.mod92	29	5702.117

attrition.mod31	5	5938.827	attrition.mod62	11	5819.63	attrition.mod93	30	5683.129

Table 19: Model selection based on BIC

	df	BIC		df	BIC		df	BIC
attrition.mod1	3	6012.922	attrition.mod32	16	6014.351	attrition.mod63	11	5908.098
attrition.mod2	4	6000.102	attrition.mod33	7	5995.214	attrition.mod64	22	5900.817
attrition.mod3	6	6005.253	attrition.mod34	7	5989.387	attrition.mod65	13	5925.675
attrition.mod4	6	5939.442	attrition.mod35	16	5983.72	attrition.mod66	13	-15086.1
attrition.mod5	4	5977.162	attrition.mod36	7	6025.538	attrition.mod67	11	5915.437
attrition.mod6	4	6005.596	attrition.mod37	7	6021.76	attrition.mod68	22	5884.352
attrition.mod7	15	6003.122	attrition.mod38	18	6023.1	attrition.mod69	13	5928.699
attrition.mod8	6	6033.834	attrition.mod39	18	6002.028	attrition.mod70	13	5924.935
attrition.mod9	6	6027.833	attrition.mod40	18	6023.1	attrition.mod71	20	5864.263
attrition.mod10	4	6018.394	attrition.mod41	18	6002.028	attrition.mod72	11	5925.758
attrition.mod11	7	5988.355	attrition.mod42	10	5915.141	attrition.mod73	11	5919.926
attrition.mod12	7	5912.896	attrition.mod43	8	5953.585	attrition.mod74	20	5964.795
attrition.mod13	5	5961.262	attrition.mod44	8	5974.688	attrition.mod75	20	5942.485
attrition.mod14	5	5993.208	attrition.mod45	19	6002.544	attrition.mod76	12	5879.509
attrition.mod15	16	6001.1	attrition.mod46	10	6010.208	attrition.mod77	23	5862.071
attrition.mod16	7	6021.929	attrition.mod47	10	6004.857	attrition.mod78	14	5899.239
attrition.mod17	7	6017.276	attrition.mod48	10	5917.953	attrition.mod79	14	5898.028
attrition.mod18	7	6017.276	attrition.mod49	10	5939.32	attrition.mod80	23	5874.639
attrition.mod19	5	6005.016	attrition.mod50	21	5917.446	attrition.mod81	14	5926.396
attrition.mod20	9	5943.515	attrition.mod51	12	5951.321	attrition.mod82	14	5923.198
attrition.mod21	7	5974.067	attrition.mod52	12	5955.325	attrition.mod83	23	5881.391
attrition.mod22	7	5992.733	attrition.mod53	8	5910.194	attrition.mod84	23	5849.319
attrition.mod23	18	6005.969	attrition.mod54	19	5872.522	attrition.mod85	24	5848.149
attrition.mod24	9	6024.816	attrition.mod55	10	5925.503	attrition.mod86	15	5893.966
attrition.mod25	9	6018.03	attrition.mod56	10	5918.831	attrition.mod87	15	5893.634
attrition.mod26	7	5909.891	attrition.mod57	17	5949.596	attrition.mod88	26	5888.877
attrition.mod27	7	5939.388	attrition.mod58	8	5985.594	attrition.mod89	26	5861.643
attrition.mod28	18	5905.221	attrition.mod59	8	5982.164	attrition.mod90	27	5867.64

attrition.mod29	9	5953.886	attrition.mod60	19	6002.544	attrition.mod91	27	5844.837
attrition.mod30	9	5952.373	attrition.mod61	19	5985.605	attrition.mod92	29	5876.301
attrition.mod31	5	5968.859	attrition.mod62	11	5885.7	attrition.mod93	30	5863.32

Saturated

attrition.mod93

Frailty distribution: gamma

Baseline hazard distribution: Weibull

Loglikelihood: -2811.565

	ESTIMATE	SE	p-val	
theta	0.215	0.103		
rho	2.423	0.075		
lambda	0.000	0.000		
Gendermale	0.437	0.096	0.000	***
ResidenceLari	-0.248	0.156	0.111	
ResidenceLimuru	0.311	0.127	0.014	*
ResidenceNairobi	0.242	0.107	0.023	*
EducationlevelDiploma	1.121	0.116	0.000	***
EducationlevelMSc	-0.001	0.119	0.994	
EducationlevelPhD	-0.040	0.145	0.780	
Age	-0.032	0.004	0.000	***
MaritalstatusSingle	0.386	0.087	0.000	***
TeachingsubjectsBiology	1.451	0.278	0.000	***
TeachingsubjectsBusiness_studies	1.099	0.275	0.000	***
TeachingsubjectsChemistry	0.270	0.271	0.318	
TeachingsubjectsComputer_Studies	1.242	0.232	0.000	***
TeachingsubjectsCRE	1.776	0.244	0.000	***
TeachingsubjectsEnglish	0.703	0.238	0.003	**
TeachingsubjectsGeography	0.986	0.234	0.000	***

```
TeachingsubjectsHistory         -0.333   0.300 0.267
TeachingsubjectsHome_science     0.708   0.398 0.075 .
TeachingsubjectsKiswahili        0.454   0.252 0.072 .
TeachingsubjectsMathematics      0.797   0.229 0.001 ***
TeachingsubjectsPhysics          0.745   0.236 0.002 **
SchoolcategoryCounty             0.271   0.132 0.040 *
SchoolcategoryExtra_county       0.079   0.126 0.530
SchoolcategoryNational           0.254   0.174 0.145
PresentassignmentHoD             0.682   0.144 0.000 ***
PresentassignmentPrincipal       0.054   0.112 0.633
PresentassignmentTeacher        -0.114   0.132 0.387
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Kendall's Tau: 0.097
```