AL KUT JOURNAL OF ECONOMIC AND ADMINISTRATIVE SCIENCES
Publisher: College of Economics and Management - Wasit University

Using The Generalized Gamma Model To Estimate The Survival Function

Rasool Hadi Abdel Moneim (1) Entsar Arebe Fadam(2)

Department of Statistics, College of Administration and Economics, University of Baghdad

rassoul.abd2101p@coadec.uobaghdad.edu.iq Entsar.arebe@coadec.uobaghdad.edu.iq

Abstract

In many medical studies, the outcome of interest is the time remaining for an event to occur. These include deaths, disease progression, or hospitalization. To aid in decision making, the hazard function is estimated from parametric models which aim to quantify the benefits to patients. Time-to-event (TTE) data with complete follow-up are rarely available. As survival function data, whether complete data or monitoring data of various types, is often in the form of a time series represented by a time period of survival for each patient or for each machine, we notice when analyzing survival data that it suffers from problems of instability or high fluctuation as a result of the abnormal phenomena of this. The idea in the research process was to process survival function data through generalized linear models by using one of the methods of generalized linear models (GLMs), represented by the generalized gamma model, to analyze time-to-event (TTE) data, through a link function linked to people at risk. At the time t that the patient is exposed to, using two methods, maximum likelihood and weighted iterative maximum potential, to predict survival data.

Keywords: survival time, hazard function, generalized linear models, maximum likelihood, iterative weighted maximum likelihood , gamma model.

1. Introduction

Survival analysis is a branch of statistics to analyze the expected time period until one or more events occur, as the response variable is the time until the event occurs, and it is often called the failure time, event time, or survival time, as in the case Medical experiences such as time to death, and time Until pregnancy, and the time until an ECG abnormality occurs during exercise on a treadmill. The interest of many studies in survival data has increased as a result of the important role that this topic plays in dealing with the timing of living organisms, such as the probability of survival and average life. Survival analysis is considered one of the important branches of statistics that has a fundamental role in analyzing any phenomenon based on the statistical data available about that phenomenon, and therefore when carrying out the estimation process for the purpose of obtaining new estimators that have the desired characteristics Find it in the ideal destination that can be relied upon to arrive To obtain more accurate results, the appropriate estimation method must be chosen as a result of the limitations imposed on standard models, as they all impose a linear assumption. Therefore, there was a need for more flexible parametric models that mitigate this assumption. In this paper, we present generalized linear models (GLMs) and show that they can express survival models and provide good results. Some of the models within the framework of GLMs were taken and used for TTE analysis, which shows that GLM-based models can provide superior within-sample estimates and more reasonable extrapolations than standard survival models

2. Reference review

Due to the importance of the survival function in the process of predicting the survival time and occurrence of an event and in various specializations, especially in the medical field, many researches have focused on this field. Some researchers mentioned the following. (Gensheimer, M.F., and Narasimhan, B., 2019) described a discrete-time survival model designed using neural networks, and they called it NNet-survival. The model was trained using the maximum likelihood method and using mini-batch stochastic gradient descent. Gradient Descent (SGD), it was found that the method used allows convergence and rapid application to large data sets, as the method was considered one of the types of deep learning. In this study, the SGD method was compared with other methods, such as Cox-nnet and Deep-surv, and the SGD method was better than others(2). (Wen, C., and Chen, Y., 2020) Initially, they studied Survival was traditionally analyzed on a continuous time scale after which observations were recorded at interrupted time intervals. The problem of missing and not measuring some information related to time-dependent covariates was addressed. The sufficient discrete hazard function (SDH) was used to address this problem. SDH is based on the conditional score idea of the logistic hazard model to deal with common variables that were not measured. In addition, the least squares method was used. To estimate the value of the missing covariate using the basics of regression, these methods were developed to deal with a single event and with competing risks[16]Lamiaa &, Entsar(2021) conditions of use was analyzed mathematically by applying structural stability analysis. As a result, we found that the probability of structural failure is an indicator that can be used to estimate whether an experiment will be accepted or rejecte[6].Lamia & Entsar (2022)The concept of structural reliability analysis presents a methodology used to verify the efficiency of engineering structures during the design and testing stages in order to support a more balanced design between the safety of the structure and its requirements Did. This is done by calculating the probability of structural failure. This is to measure how well the design under study performs depending on the drag coefficient (component) and strength it depends on. In this study, the researcher proposes that he develop a third-order moment. When applied to numerical examples, the method using the Downhill Simplex algorithm was highly efficient compared to his original Third Moment method.In this study, we analyzed the structural reliability to ensure the strength of the structural design and conducted experiments to confirm the effect of G.U.M mouthwash on vision of curing and rehealing. Composite dental fillings are light-cured (VLC). Experiments were carried out in accordance with agreed international standards and used the analytical technique "Copula D-vine" to determine the possibility of structural failure when dental fillings are exposed to G.U.M mouthwash[6]. Noor & Entsar (2023) These images were incorporated into a machine learning method to estimate patient survival. This relies on his two feature extraction methods. One is a fast independent component algorithm (fast ICA algorithm) and the other is nonnegative matrix factorization (NMF algorithm). It uses two machine learning algorithms. The first is the Random Survival Forest algorithm and the second is the Support Vector Machine (SVM) algorithm. By applying supervised machine learning techniques to mammograms of breast cancer patients, the resulting model is optimal for estimating survival based on mean squared error (MSE) and concordance index (C-Index). I found out something. We compared the results using the Fast ICA and Random Survival Forest algorithms with three other her models. Therefore, we encourage interested healthcare institutions and institutions to adopt this model[13].

3. Concept of Survival Analysis[5][6][10][14]

Interest in survivability analysis began about half a century ago with the beginning of conducting engineering applications during World War II on the reliability or reliability of military equipment (estimating the reliability of military equipment). This interest continued into the post-war period and expanded to include military, industrial, and commercial products. Most statistical research has focused on Study and analysis of parametric survival models to represent data of positive random variables that have specific distributions according to the nature of survival times for the phenomena studied, as the random variable that we are interested in studying represents the time until the occurrence of the event that often follows one of the asymmetric probability distributions. The exponential family distributions (exponential, kamma) are considered , beta, Weibull, and the natural logarithm distribution (one of the most widely used distributions in survival analysis. The dependent variable in regression models of survival data is time measured in years, months, weeks, or days. The entry of the individual into the experiment is measured from the beginning of follow-up of the individual until the occurrence of the end event (death). Or recovery (and in medical research, the start time is often from the date of registration of the case or individual in a particular study until the end of the study time, as survival analysis includes a group of statistical methods (parametric and non-parametric) used in analyzing random variables with positive values only, and within the past two decades, The number of clinical trials in the field of medical research increased to focus on modern nonparametric methods and their applications in survival analysis, while most previous statistical research for engineering applications focused on studying and estimating parametric survival models.

4. Survival Data [10][15]

The time values for each unit are recorded within the study or experiment from the beginning of the study or experiment until the occurrence of the event under study (time-to-event) or the end of the study time measured in one of the units of time (minute, hour, day, etc.). (It is called survival data or survival times. Accordingly, survival data is described by a random variable (T) whose domain is the set of positive values. Accordingly, it can be represented by one of the probability distributions with positive values whose probability function is asymmetrical and skewed to the right (positive skewness). Therefore, it cannot be assumed that the data Survival follows a symmetrical normal distribution, and therefore statistical methods based on the assumption of a normal distribution cannot be applied to survival data. Hence, researchers began more than half a century ago to search and investigate alternative statistical methods. During the journey of research and study in the field of survival data analysis, researchers faced a challenge. Of another type, there is a possibility that some of the cases included in the study cannot be observed for their survival time because the event under study does not happen to the individual before the end of the study time. Or some subjects leave the study or are lost before the end of the study time. Survival data has been classified into (Complete Survival Data, Incomplete Survival Data, and Censoring Data). It is of many types, including (Right Censoring Data [7,13] and Left Censoring Data3) Interval Censoring

5. General Linear Models

The General Linear Model, or the so-called regression model, can be expressed as the Multiple Linear Regression Model, which contains one dependent variable and several independent (explanatory) variables, as follows (Yan & Gang Su, 2009) and (Madsen & Thyregod). , 2010)

$$\underline{y} = X\underline{\beta} + \underline{\varepsilon} \qquad \ldots \ldots (1)$$

Since:

y: vector of observations of the dependent variable with dimension (n×1), where n represents the number of observations.

χ: An information matrix representing the observations of the explanatory variables with the dimension (n×p), where p-1 represents the number of explanatory variables

β: vector of regression parameters with dimension (p×1), which is often called the fixed effect.

ε: vector of random errors with dimension (n×1), which follows a multivariate normal distribution.

5-1. Generalized Linear Models (4)(8)(9)

The researchers (Nelder & Wedderburn, 1972) are the first to introduce generalized linear models, and then they were developed before (Nelder & McCulloch, 1989). Generalized linear regression models (GLMS) are a generalization of general linear regression models, and this generalization includes probability distributions that It belongs to the exponential family through two aspects: the first is that the errors do not follow a normal distribution, that is, they follow one of the exponential distributions, and the second is that the average response is not necessarily taken as a linear set of parameters and explanatory variables, but rather it is a linear function with a middle called the line function, that is, it is Modeling a function of the mean instead of the mean itself. As for the variance of the observations, it is a function of the mean (). A generalized linear model (GLM) is formed on the following components:

The random component is the response variable that has a probability distribution that belongs to the exponential family. The Linear Systematic Component, which contains the observations of the explanatory variables (X) and the vector of parameters (β) contained in the general linear model equation, as the systematic part is $\eta_i = \underline{x}_i^T \underline{\beta}$ which is called the Linear Predictor. It is given in the following formula:

$$\eta_i = \underline{x}_i^T \underline{\beta} = \beta_0 + \beta_1 x_{i1} + \ldots + \beta_{(p-1)} x_{i(p-1)} \quad \ldots\ldots\ldots (2)$$

Link Function

The link function g(.) represents the link between the random component and the regular component, which determines the linear relationship between the average of the response variable (y) and the linear predictor. It is a differentiable function that links the random part and the regular part, as the inverse of the link function at the regular part is equal to the expectation of the response variable

$$E(y_i) = g^{-1}(\eta_i) = \mu_i$$

$$\eta_i = g(\mu_i) = x_i^T \beta$$

Where the link function of the distribution is obtained from its exponential family formula from the legal parameter θ accompanying the response variable, it follows:

$$g(\mu_i) = \theta_i = \eta_i$$

Whereas in the general linear model, the observations are independent, the average of the observations is a linear function in terms of the independent variables, and the variance of the observations is constant. In the generalized linear model, the independence condition is maintained, while the mean condition becomes more flexible, such that it is a linear function in terms of the independent variables via a link function, and the variance of observations is a function in terms of the mean, meaning that (Hardin & Hilbe, 2018)

$$y_i \sim \text{indp. EF}(\mu_i, h(\mu_i))$$

Since $h(\mu_i)$ is a function of the mean μ_i.

All distributions that belong to the exponential family can be expressed in a unified form, and based on that, the Likelihood Function for any random variable belonging to the exponential family can be written as follows (Brown & Prescott, 2015):

$$L\left(\underline{\theta};\underline{y}\right) = \exp\left[\sum_{i=1}^{n}\left(\frac{y_i\theta_i - b(\theta_i)}{a_i(\phi)}\right) + C\right] \quad \dots\dots(3)$$

Since C is a constant, θi is replaced by a linear predictor. One of the advantages of the generalized linear model ($\boldsymbol{GLM_Z}$) It frees the data from the condition of linearity in the relationship $\mathbf{Y}$ with $\mathbf{X}_s$ and also frees researchers from the analysis restrictions imposed on the regression function. Thus, many distributions can be modeled {normal distribution, binomial distribution, Poisson distribution, inverse normal, inverse binomial, gamma distribution... ..}

When using GLM models for event data over a period of time, yt is associated with the people at risk at time t

yt ~ exponential family distribution

Observation model: E[yt] = $\eta_t \ \tau_t$

Response function: η_t = h(x_t^Tß$_t$)

ß: is a vector of parameter coe_cients to be estimated from the data,

xt : is a covariate, assumed known (with transpose xt T)),

Where $\tau_t$represents information about the sample size "at risk" in the time period (t_i; t_i+1) and is calculated by the formula:-

τ_t = (t_{i+1}- t_i)(n_{ti} –c_{ti} /2) ……(4)

whereas

n_{ti}: the number of people at risk at time t

c_{ti}: number of people subject to supervision in the period (ti; ti+1)

5-1-1-.Gamma regression model(4)(12)(17)

The Gamma Model (GRM) is used to model positive skewed continuous data such as survival data, and here the response variable follows the Gamma distribution, and it can be formulated according to the general form of the exponential family as follows ((Faraway, 2016), (Hardin & Hilbe, 2018), and (McCullagh & Nelder) , 1989):

$$f(y_i; \alpha_i, \mu_i) = \frac{1}{\Gamma(\alpha_i)} \left(\frac{\alpha_i}{\mu_i}\right)^{\alpha_i} y_i^{\alpha_i - 1} e^{-\frac{\alpha_i y_i}{\mu_i}} \qquad \ldots\ldots(5)$$

$$\Rightarrow f(y_i; \alpha_i, \mu_i) = exp\left[-\frac{\alpha_i y_i}{\mu_i} - \alpha_i \, ln(\mu_i) + \alpha_i \, ln(\alpha_i) + (\alpha_i - 1)\, ln(y_i) - ln\big(\Gamma(\alpha_i)\big)\right]$$

$$\Rightarrow f(y_i; \alpha_i, \mu_i) = exp\left[\frac{y_i/\mu_i + ln(\mu_i)}{-1/\alpha_i} + \alpha_i \, ln(\alpha_i) + (\alpha_i - 1)\, ln(y_i) - ln\big(\Gamma(\alpha_i)\big)\right]$$

We note that the probability density function of the gamma distribution becomes similar to the general formula of the exponential family, as:

$$\theta_i = \frac{1}{\mu_i}$$

$$b(\theta_i) = -\, ln(\mu_i)$$

$$a(\phi) = -\, 1/\alpha_i$$

$$c(y_i, \phi) = \alpha_i \, ln(\alpha_i) + (\alpha_i - 1)\, ln(y_i) - ln\big(\Gamma(\alpha_i)\big)$$

As for expectation and variance, the function $b(\theta_i)$ must be found in terms of θ_i, as follows:

$$\theta_i = \frac{1}{\mu_i}$$

$$\Rightarrow \mu_i = \frac{1}{\theta_i}$$

Now by substituting μ_i into the equation $b(\theta_i)$ we get:

$$b(\theta_i) = -\, ln\left(\frac{1}{\theta_i}\right)$$

Taking the first derivative of the function $b(\theta_i)$ relative to θ_i we get:

$$b'(\theta_i) = \frac{1}{\theta_i}$$

Also, the second derivative of the function $b(\theta_i)$relative to θ_i is:

$$b''(\theta_i) = \frac{-1}{\theta_i^2}$$

Now by substituting we get both the expectation and the variance:

$$E\left(y_i|\underline{x}_i^T,\underline{\beta}\right) = \mu_i = b'(\theta_i) = \frac{1}{\theta_i} = \mu_i$$

$$Var\left(y_i|\underline{x}_i^T,\underline{\beta}\right) = b''(\theta_i)a(\phi) = \frac{-1}{\theta_i^2}(-1/\alpha_i) = \frac{1}{\alpha_i\theta_i^2} = \frac{\mu_i^2}{\alpha_i}$$

$$g(\mu_i) = \theta_i = \frac{1}{\mu_i} = \mu_i^{-1}$$

The potential function is obtained in the case of a gamma distribution, as follows

$$L\left(\underline{\mu};\underline{y}\right) = exp\left[\sum_{i=1}^{n}\left(\frac{y_i/\mu_i\tau_t + ln(\mu_i\tau_t)}{-1/\lambda} + \lambda\{ln(\lambda) + ln(y_i)\} - ln\big(\Gamma(\lambda)\big) - ln\{y_i\}\right)\right] \quad \ldots(6)$$

Substituting the inverse of the link function, we obtain

$$L\left(\underline{\beta};\underline{y}\right) = exp\left[\sum_{i=1}^{n}\left(\frac{y_i\eta_i - ln(\eta_i)}{-1/\lambda} + \lambda\{ln(\lambda) + ln(y_i)\} - ln\big(\Gamma(\lambda)\big) - ln\{y_i\}\right)\right] \quad \ldots(7)$$

After replacing the linear predictor, we obtain the potential function for the generalized linear model (GLM) in the case of a gamma distribution

$$L\left(\underline{\beta};\underline{y}\right) = exp\left[\sum_{i=1}^{n}\left(\frac{y_i\left(\underline{x}_i^T\underline{\beta}\right) - ln\left(\underline{x}_i^T\underline{\beta}\right)}{-1/\lambda} + \lambda\{ln(\lambda) + ln(y_i)\} - ln\big(\Gamma(\lambda)\big) - ln\{y_i\}\right)\right] \ldots.(8)$$

6. **Estimation methods**:-

6-1.The maximum likelihood method

$$f(a,\lambda,x) = \frac{\lambda^a}{\Gamma(a)} \; x_i^{a-1} \quad e^{-\lambda x_i}$$

Suppose Xi ~ Gamma(a, λ) and i.i.d.

$$l(a,\lambda,x_i) = \prod_{i=1}^{n} \frac{\lambda^a}{\Gamma(a)} \; x_i^{a-1} \quad e^{-\lambda x_i}$$

$$l(a,\lambda,) = \sum_{i=1}^{n} log \frac{\lambda^a}{\Gamma(a)} \; x_i^{a-1} \quad e^{-\lambda x_i}$$

$$l(a,\lambda,) = \sum_{i=1}^{n} \log(\lambda^a) - \sum_{i=1}^{n} \log(\Gamma(a)) + (\mathrm{a}-1)\sum_{i=1}^{n} \log(x_i) - \lambda \sum_{i=1}^{n} x_i$$

$$l(a,\lambda,) = na\log(\lambda) - \mathrm{n}\log(\Gamma(a)) + (\mathrm{a}-1)\sum_{i=1}^{n} \log(x_i) - \lambda \sum_{i=1}^{n} x_i$$

And set to zero (solve the λ-partial first):

$$\frac{\partial l(a,\lambda)}{\partial \lambda} = \frac{na}{\hat{\lambda}} - \sum_{i=1}^{n} x_i = 0$$

$$\hat{\lambda} = \frac{\hat{a}}{n^{-1}\sum_{i=1}^{n} x_i} = \frac{\hat{a}}{\bar{x}}$$

And plug into a-partial equation

$$\frac{\partial l(a,\lambda)}{\partial a} = \mathrm{n}\log(\lambda) - \mathrm{n}\frac{\bar{\Gamma}(\hat{a})}{\Gamma(\hat{a})} + \sum_{i=1}^{n} x_i = 0$$

$$\frac{\partial l(a,\lambda)}{\partial a} = \log(\hat{a}) - \log(\bar{x}) - \frac{\bar{\Gamma}(\hat{a})}{\Gamma(\hat{a})} + \frac{1}{n}\sum_{i=1}^{n} x_i = 0$$

For a particular application, we need a numerical method to solve. What are some examples? bisection method Newton's method (also called Newton-Raphson)

6-2. Iterative weighted maximum likelihood method (26)

The parameters of the Generalized Linear Regression Model (GLM) are estimated, as indicated by researchers Nelder-Weddrburn, that the majority of cases that can be modeled with the $\mathrm{GLM_Z}$ method will follow the same method in estimating their parameters using the maximum likelihood method, and that estimating the model parameters will depend on (linearization) . The link function $\eta = G(X^T\beta)$will be the focus of interest instead of β.

Since:- $G(yi) \approx G(\mu) + (y_i - \mu)G'(\mu)$ $G'(\mu) = \frac{d\eta_i}{d\mu_i}$

It is the derivative of the link function with respect to μ

Substituting $G(y_i)=\hat{Z}_i$, $G(\mu_i)= \hat{\eta}_i$ that $\hat{\eta}_i$represents the estimate of the link function, and the formula specified by Nelder-McCullagh within the iterative weighted least squares (IWLS) method will be:

$$\hat{\beta}^{new} = \hat{\beta}^{old} + (X^T W X)^{-1} X^T (\underline{y} - \underline{\mu}) \ldots\ldots\ldots\ldots(11)$$

Since:- $W = diag[\frac{1}{b''(\theta_i)(\frac{d\eta_i}{d\mu_i})^2}]$

$$\mu^T = [n_1\theta_1 \ n_2\theta_2 \ \ldots\ldots\ldots.. n_i\theta_i]$$

$$\hat{Z}i = \hat{\eta} + (y_i - \mu_i)\frac{d\hat{\eta}_i}{d\mu_i} \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(12)$$

The formula $\hat{\boldsymbol{\beta}}^{\boldsymbol{new}}$ can be described by the following equation:

$$\hat{\beta}^{new} = (X^T W X)^{-1} X^T W Z \qquad \ldots\ldots\ldots\ldots\ldots\ldots..(13)$$

The estimator $\hat{\boldsymbol{\beta}}^{\boldsymbol{new}}$ asymptotically, contains all the good characteristics of OLS, as:

$$\beta_{IWLS} \sim N(\beta , \sigma^2 (X^T W X)^{-1})$$

7. Test Dat

After the process of collecting and tabulating the data, we move to the stage of testing the data regarding the quality of fit (goodness of fit) for the distribution of the response variable shown in Table No. (1) using the ready-made statistical program (SPSS). Accordingly, the Kolmogorov-Smirnov test was applied according to the following hypothesis:

$$H_0: Y \sim Gamma(\mu)$$

$$H_1: Y \nsim Gamma(\mu)$$

Table (1) test results Kolmogorov –Smirnov

Var 1	Statistic	Observation	Sig. (p. value)
	0.06506	200	0.35052

From Table (1), it is noted that the probability value (P-value) of the Kolmogorov-Smirnov test is equal to (0.35052), which is greater than the value of the test statistic at the significance level (0.05). Therefore, the null hypothesis which states that the data follows a Gamma distribution is not rejected, and the figure is The following shows the Gamma distribution curve:

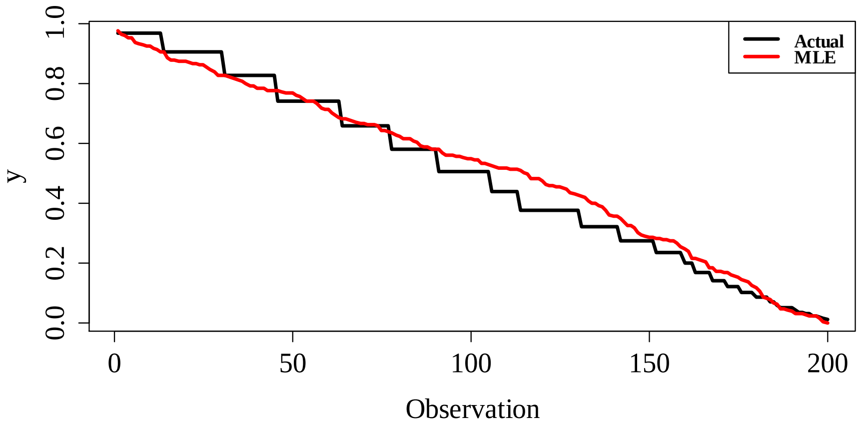
9. Model estimation

Two methods, maximum likelihood and weighted iterative maximum likelihood, were used to estimate the regression parameters, and based on that, the estimated Table (3)

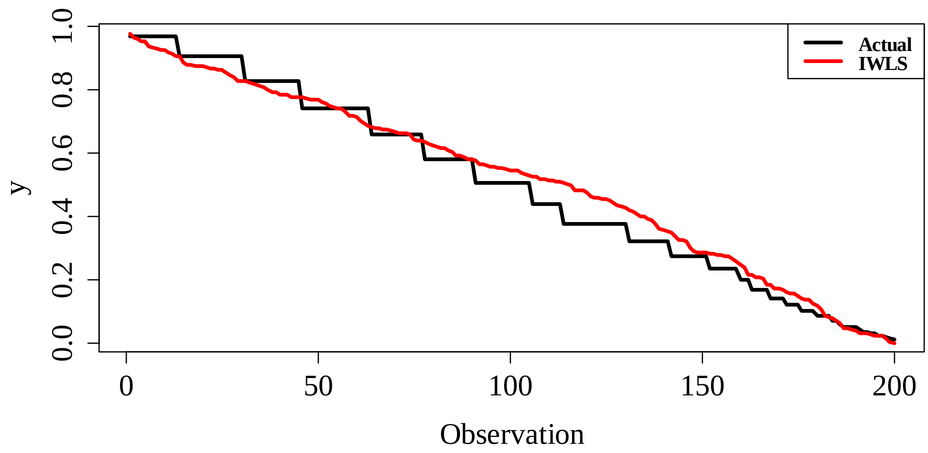
Table (3) Parameter estimates by the maximum likelihood method and the weighted iterative maximum likelihood method

Methods	MLE	IWLS
β_0	2.19544	2.42636
β_1	0.91599	0.92853
β_2	-0.56767	-0.60549
β_3	-0.02130	-0.02144
β_4	-0.16820	-0.16328
β_5	-0.25546	-0.23648
β_6	-0.20861	-0.21250
β_7	2.80708	2.82274
β_8	-0.07747	-0.07410
MSE	0.18755	0.18265

We note from Table (3) that the values of the parameters represent the logical form of the variables with respect to their effect on myocardial infarction, and that the iterative weighted maximum potential method is considered the best way to estimate the survival function, as the value of (MSE) is (0.18265), which is the smallest value and with a slight difference from the potential. The greatest



Figure(1) showing the survival function according to the method (MLE)



Figure(2) showing the survival functionaccording to the method (IWLS)

Table (4) shows the survival function for (R, MLE, IWLS)

S(IWLS)	S(MLE)	S(R)	TIME	S(IWLS)	S(MLE)	S(R)	TIME	
0.8291	0.8295	0.8281	0.4286	0.9770	0.9770	0.9701	0.1429	1
0.8261	0.8252	0.8281	0.4286	0.9668	0.9664	0.9701	0.1429	2
0.8213	0.8199	0.8281	0.4286	0.9606	0.9604	0.9701	0.1429	3
0.8155	0.8147	0.8281	0.4286	0.9558	0.9554	0.9701	0.1429	4
0.8110	0.8134	0.8281	0.4286	0.9549	0.9547	0.9701	0.1429	5
0.8103	0.8096	0.8281	0.4286	0.9376	0.9375	0.9701	0.1429	6
0.8013	0.8013	0.8281	0.4286	0.9350	0.9347	0.9701	0.1429	7
0.7947	0.7930	0.8281	0.4286	0.9304	0.9305	0.9701	0.1429	8
0.7915	0.7917	0.8281	0.4286	0.9278	0.9286	0.9701	0.1429	9
0.7854	0.7856	0.8281	0.4286	0.9255	0.9248	0.9701	0.1429	10
0.7853	0.7847	0.8281	0.4286	0.9183	0.9183	0.9701	0.1429	11
0.7838	0.7842	0.8281	0.4286	0.9156	0.9150	0.9701	0.1429	12
0.7787	0.7781	0.8281	0.4286	0.9083	0.9082	0.9701	0.1429	13
0.7770	0.7775	0.8281	0.4286	0.9065	0.9064	0.9068	0.2857	14
0.7757	0.7769	0.8281	0.4286	0.8877	0.8877	0.9068	0.2857	15
0.7753	0.7765	0.7438	0.5714	0.8806	0.8802	0.9068	0.2857	16
0.7722	0.7723	0.7438	0.5714	0.8790	0.8789	0.9068	0.2857	17
0.7702	0.7711	0.7438	0.5714	0.8767	0.8765	0.9068	0.2857	18
0.7692	0.7709	0.7438	0.5714	0.8750	0.8747	0.9068	0.2857	19
0.7675	0.7687	0.7438	0.5714	0.8750	0.8743	0.9068	0.2857	20

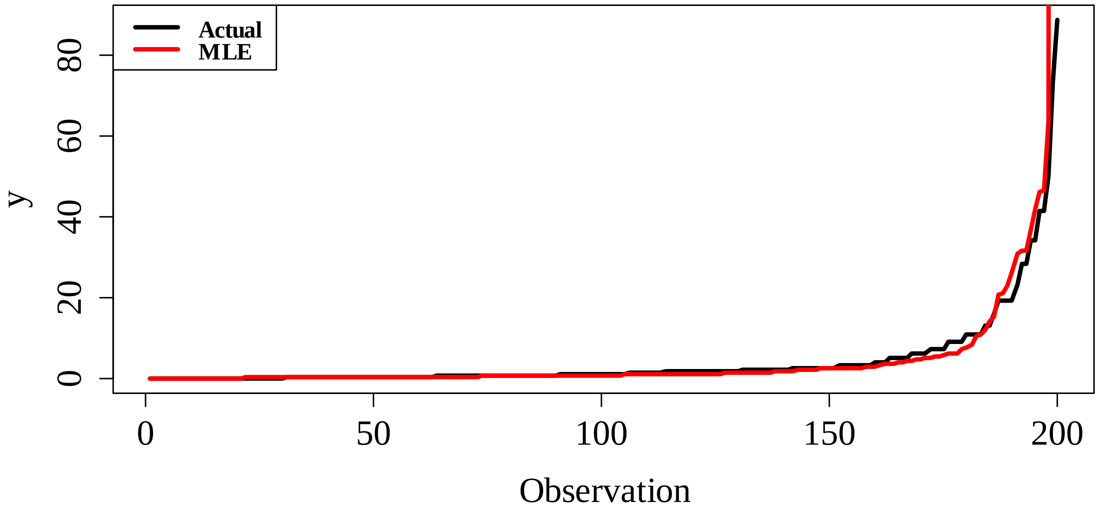
0.7615	0.7612	0.7438	0.5714	0.8707	0.8713	0.9068	0.2857	21
0.7594	0.7588	0.7438	0.5714	0.8681	0.8674	0.9068	0.2857	22
0.7517	0.7515	0.7438	0.5714	0.8673	0.8663	0.9068	0.2857	23
0.7439	0.7434	0.7438	0.5714	0.8651	0.8649	0.9068	0.2857	24
0.7421	0.7431	0.7438	0.5714	0.8633	0.8633	0.9068	0.2857	25
0.7416	0.7418	0.7438	0.5714	0.8561	0.8550	0.9068	0.2857	26
0.7351	0.7345	0.7438	0.5714	0.8473	0.8464	0.9068	0.2857	27
0.7202	0.7194	0.7438	0.5714	0.8384	0.8383	0.9068	0.2857	28
0.7168	0.7159	0.7438	0.5714	0.8303	0.8301	0.9068	0.2857	29
0.7143	0.7125	0.7438	0.5714	0.8303	0.8299	0.9068	0.2857	30
0.5778	0.5787	0.5051	1.0000	0.7040	0.7043	0.7438	0.5714	31
0.5654	0.5680	0.5051	1.0000	0.6944	0.6940	0.7438	0.5714	32
0.5642	0.5623	0.5051	1.0000	0.6866	0.6879	0.7438	0.5714	33
0.5606	0.5619	0.5051	1.0000	0.6839	0.6838	0.6598	0.7143	34
0.5582	0.5591	0.5051	1.0000	0.6802	0.6814	0.6598	0.7143	35
0.5580	0.5586	0.5051	1.0000	0.6800	0.6802	0.6598	0.7143	36
0.5546	0.5556	0.5051	1.0000	0.6755	0.6758	0.6598	0.7143	37
0.5522	0.5532	0.5051	1.0000	0.6734	0.6717	0.6598	0.7143	38
0.5505	0.5510	0.5051	1.0000	0.6702	0.6662	0.6598	0.7143	39
0.5464	0.5473	0.5051	1.0000	0.6681	0.6660	0.6598	0.7143	40
0.5449	0.5460	0.5051	1.0000	0.6652	0.6649	0.6598	0.7143	41
0.5436	0.5453	0.5051	1.0000	0.6622	0.6629	0.6598	0.7143	42

0.5356	0.5344	0.5051	1.0000	0.6614	0.6619	0.6598	0.7143	43
0.5339	0.5343	0.5051	1.0000	0.6593	0.6575	0.6598	0.7143	44
0.5278	0.5292	0.5051	1.0000	0.6453	0.6428	0.6598	0.7143	45
0.5258	0.5240	0.4374	1.1429	0.6413	0.6417	0.6598	0.7143	46
0.5256	0.5229	0.4374	1.1429	0.6393	0.6402	0.6598	0.7143	47
0.5183	0.5173	0.4374	1.1429	0.6358	0.6364	0.5796	0.8571	48
0.5168	0.5172	0.4374	1.1429	0.6295	0.6280	0.5796	0.8571	49
0.5148	0.5161	0.4374	1.1429	0.6227	0.6232	0.5796	0.8571	50
0.5138	0.5141	0.4374	1.1429	0.6181	0.6173	0.5796	0.8571	51
0.5118	0.5125	0.4374	1.1429	0.6171	0.6167	0.5796	0.8571	52
0.5107	0.5120	0.4374	1.1429	0.6163	0.6152	0.5796	0.8571	53
0.5075	0.5082	0.3766	1.2857	0.6095	0.6087	0.5796	0.8571	54
0.5007	0.5004	0.3766	1.2857	0.6025	0.6038	0.5796	0.8571	55
0.4969	0.4971	0.3766	1.2857	0.5934	0.5910	0.5796	0.8571	56
0.4843	0.4837	0.3766	1.2857	0.5911	0.5903	0.5796	0.8571	57
0.4826	0.4829	0.3766	1.2857	0.5900	0.5893	0.5796	0.8571	58
0.4819	0.4813	0.3766	1.2857	0.5814	0.5803	0.5796	0.8571	59
0.4726	0.4733	0.3766	1.2857	0.5802	0.5793	0.5796	0.8571	60
0.2841	0.2856	0.2755	1.5714	0.4611	0.4621	0.3766	1.2857	61
0.2835	0.2825	0.2344	1.7143	0.4591	0.4600	0.3766	1.2857	62
0.2808	0.2822	0.2344	1.7143	0.4576	0.4590	0.3766	1.2857	63
0.2778	0.2789	0.2344	1.7143	0.4539	0.4545	0.3766	1.2857	64

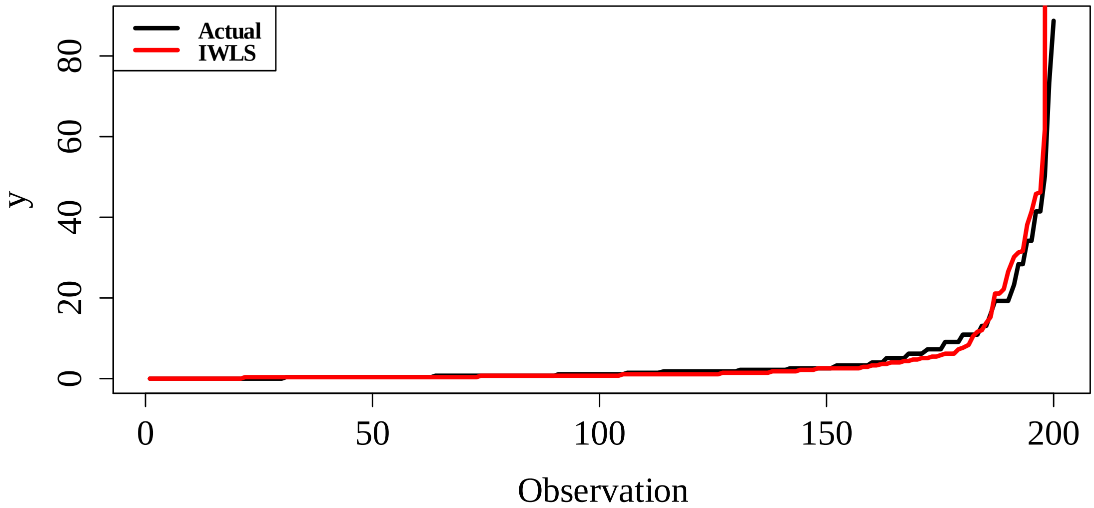
0.2772	0.2774	0.2344	1.7143	0.4537	0.4536	0.3766	1.2857	65
0.2758	0.2734	0.2344	1.7143	0.4495	0.4520	0.3766	1.2857	66
0.2738	0.2727	0.2344	1.7143	0.4438	0.4453	0.3766	1.2857	67
0.2659	0.2671	0.2344	1.7143	0.4339	0.4335	0.3766	1.2857	68
0.2565	0.2551	0.2344	1.7143	0.4318	0.4313	0.3766	1.2857	69
0.2459	0.2464	0.1987	1.8571	0.4278	0.4282	0.3766	1.2857	70
0.2403	0.2401	0.1987	1.8571	0.4211	0.4227	0.3228	1.4286	71
0.2148	0.2168	0.1987	1.8571	0.4169	0.4181	0.3228	1.4286	72
0.2143	0.2134	0.1681	2.0000	0.4060	0.4065	0.3228	1.4286	73
0.2083	0.2097	0.1681	2.0000	0.3981	0.4015	0.3228	1.4286	74
0.2068	0.2073	0.1681	2.0000	0.3980	0.3995	0.3228	1.4286	75
0.2040	0.2041	0.1681	2.0000	0.3906	0.3919	0.3228	1.4286	76
0.1851	0.1850	0.1681	2.0000	0.3888	0.3883	0.3228	1.4286	77
0.1831	0.1818	0.1418	2.1429	0.3760	0.3748	0.3228	1.4286	78
0.1716	0.1710	0.1418	2.1429	0.3596	0.3613	0.3228	1.4286	79
0.1699	0.1706	0.1418	2.1429	0.3570	0.3584	0.3228	1.4286	80
0.1695	0.1695	0.1418	2.1429	0.3526	0.3545	0.3228	1.4286	81
0.1607	0.1659	0.1193	2.2857	0.3498	0.3498	0.2755	1.5714	82
0.1567	0.1579	0.1193	2.2857	0.3374	0.3363	0.2755	1.5714	83
0.1562	0.1551	0.1193	2.2857	0.3249	0.3253	0.2755	1.5714	84
0.1485	0.1506	0.1193	2.2857	0.3235	0.3242	0.2755	1.5714	85
0.1406	0.1429	0.1002	2.4286	0.3205	0.3178	0.2755	1.5714	86

0.1376	0.1401	0.1002	2.4286	0.3003	0.3008	0.2755	1.5714	87
0.1372	0.1374	0.1002	2.4286	0.2913	0.2921	0.2755	1.5714	88
0.1229	0.1232	0.1002	2.4286	0.2854	0.2878	0.2755	1.5714	89
0.1172	0.1160	0.0841	2.5714	0.2846	0.2867	0.2755	1.5714	90
0.0319	0.0314	0.0409	3.1429	0.1051	0.1055	0.0841	2.5714	91
0.0308	0.0307	0.0341	3.2857	0.0867	0.0868	0.0841	2.5714	92
0.0307	0.0307	0.0341	3.2857	0.0801	0.0831	0.0841	2.5714	93
0.0254	0.0264	0.0283	3.4286	0.0776	0.0761	0.0704	2.7143	94
0.0235	0.0233	0.0283	3.4286	0.0675	0.0660	0.0704	2.7143	95
0.0213	0.0212	0.0236	3.5714	0.0611	0.0608	0.0588	2.8571	96
0.0211	0.0211	0.0236	3.5714	0.0454	0.0457	0.0491	3.0000	97
0.0159	0.0154	0.0195	3.7143	0.0452	0.0452	0.0491	3.0000	98
0.0009	0.0009	0.0134	4.0000	0.0428	0.0420	0.0491	3.0000	99
0.0000	0.0000	0.0111	4.1429	0.0362	0.0366	0.0491	3.0000	100

We notice from Figures (1,2) that the probability values of survival decrease with increasing time, and they are close to representing the true values. Likewise, Table (4) that the probability values of the survival function are decreasing as time increases, and this reflects that the survival function is inversely proportional to time, as the duration decreases. Patient survival increases as time increases in stagnation



Figure(3) showing the Hazard function according to the method (MLE)



Figure(4) showing the Hazard function according to the method (IWLS)

Table (5) shows the Hazard function for (R, MLE, IWLS)

h(IWLS)	h(MLE)	h(Real)	h(IWLS)	h(MLE)	h(Real)	
0.8351	0.8314	0.9797	0.0235	0.0236	0.0308	1
0.8397	0.834	0.9797	0.0343	0.0347	0.0308	2

0.8671	0.8713	0.9797	0.041	0.0412	0.0308	3
0.8732	0.8717	0.9797	0.0463	0.0466	0.0308	4
0.8945	0.8895	0.9797	0.0472	0.0475	0.0308	5
0.9018	0.9084	1.2865	0.0665	0.0667	0.0308	6
0.9025	0.9125	1.2865	0.0696	0.0699	0.0308	7
0.9292	0.933	1.2865	0.0748	0.0747	0.0308	8
0.9349	0.9335	1.2865	0.0778	0.0769	0.0308	9
0.9425	0.9377	1.2865	0.0805	0.0813	0.0308	10
0.9464	0.9453	1.2865	0.089	0.089	0.0308	11
0.9538	0.9514	1.2865	0.0921	0.0929	0.0308	12
0.958	0.9531	1.2865	0.1009	0.1011	0.0308	13
0.9705	0.9676	1.6553	0.1031	0.1033	0.1028	14
0.997	0.9986	1.6553	0.1265	0.1265	0.1028	15
1.0123	1.0116	1.6553	0.1356	0.1361	0.1028	16
1.0648	1.0674	1.6553	0.1376	0.1378	0.1028	17
1.0722	1.0708	1.6553	0.1407	0.1409	0.1028	18
1.0753	1.0776	1.6553	0.1429	0.1433	0.1028	19
1.1159	1.1127	1.6553	0.1429	0.1438	0.1028	20
1.1689	1.1642	1.6553	0.1485	0.1476	0.1028	21
1.1781	1.1739	1.6553	0.1519	0.1529	0.1028	22
1.1851	1.1785	1.6553	0.153	0.1543	0.1028	23
1.2033	1.2003	1.6553	0.1559	0.1562	0.1028	24

1.2039	1.2048	1.6553	0.1583	0.1583	0.1028	25
1.2245	1.2122	1.6553	0.1681	0.1696	0.1028	26
1.2534	1.2454	1.6553	0.1802	0.1815	0.1028	27
1.3048	1.3067	1.6553	0.1927	0.1929	0.1028	28
1.3157	1.3184	1.6553	0.2043	0.2047	0.1028	29
1.3375	1.3354	1.6553	0.2044	0.2049	0.1028	30
1.375	1.3655	2.0982	0.2061	0.2056	0.2076	31
1.3986	1.3917	2.0982	0.2104	0.2118	0.2076	32
1.4629	1.4601	2.0982	0.2176	0.2197	0.2076	33
1.5121	1.4906	2.0982	0.2262	0.2275	0.2076	34
1.5129	1.5033	2.0982	0.233	0.2293	0.2076	35
1.5601	1.5518	2.0982	0.2341	0.2352	0.2076	36
1.5722	1.5751	2.0982	0.248	0.248	0.2076	37
1.6595	1.668	2.0982	0.2584	0.261	0.2076	38
1.7805	1.7681	2.0982	0.2634	0.2631	0.2076	39
1.8012	1.7899	2.0982	0.2732	0.273	0.2076	40
1.8364	1.8205	2.0982	0.2734	0.2744	0.2076	41
1.8592	1.8586	2.6295	0.2759	0.2751	0.2076	42
1.9638	1.9735	2.6295	0.2841	0.2851	0.2076	43
2.0782	2.0739	2.6295	0.2871	0.2862	0.2076	44
2.0916	2.0848	2.6295	0.2892	0.2872	0.2076	45
2.1205	2.1464	2.6295	0.2898	0.2879	0.3444	46

2.3298	2.3244	2.6295	0.295	0.2948	0.3444	47
2.433	2.424	2.6295	0.2984	0.2969	0.3444	48
2.5044	2.4748	2.6295	0.3	0.2972	0.3444	49
2.514	2.4879	2.6295	0.3029	0.3008	0.3444	50
2.52	2.5016	2.6295	0.3132	0.3137	0.3444	51
2.5273	2.5394	3.2669	0.3168	0.3179	0.3444	52
2.5614	2.5438	3.2669	0.3304	0.3306	0.3444	53
2.6	2.5856	3.2669	0.3443	0.3452	0.3444	54
2.6077	2.6043	3.2669	0.3475	0.3457	0.3444	55
2.626	2.6576	3.2669	0.3484	0.3481	0.3444	56
2.6526	2.6666	3.2669	0.3603	0.3615	0.3444	57
2.7602	2.7434	3.2669	0.3885	0.3901	0.3444	58
2.8989	2.9194	3.2669	0.3951	0.3968	0.3444	59
3.0665	3.0581	4.0318	0.4	0.4034	0.3444	60
3.1607	3.1654	4.0318	0.4204	0.4199	0.3444	61
3.6558	3.6127	4.0318	0.4401	0.4409	0.3444	62
3.6672	3.6862	4.9503	0.4564	0.4538	0.3444	63
3.8004	3.7692	4.9503	0.4623	0.4625	0.5155	64
3.8362	3.8241	4.9503	0.4701	0.4675	0.5155	65
3.9011	3.8997	4.9503	0.4706	0.4702	0.5155	66
4.4038	4.4043	4.9503	0.4805	0.4797	0.5155	67
4.4608	4.5001	6.0536	0.4849	0.4887	0.5155	68

4.8264	4.8474	6.0536	0.4921	0.5011	0.5155	69
4.8855	4.8624	6.0536	0.4967	0.5014	0.5155	70
4.9003	4.9009	6.0536	0.5034	0.504	0.5155	71
5.2224	5.028	7.38	0.5101	0.5085	0.5155	72
5.3796	5.3319	7.38	0.5119	0.5107	0.5155	73
5.4015	5.4457	7.38	0.5168	0.5208	0.5155	74
5.7346	5.6397	7.38	0.5497	0.5557	0.5155	75
6.1141	5.998	8.9757	0.5593	0.5585	0.5155	76
6.2661	6.1353	8.9757	0.5643	0.562	0.5155	77
6.2865	6.2783	8.9757	0.5727	0.5714	0.7253	78
7.1374	7.1168	8.9757	0.5884	0.5922	0.7253	79
7.5312	7.6174	10.8967	0.606	0.6045	0.7253	80
8.5111	8.479	10.8967	0.6178	0.6199	0.7253	81
10.5329	10.52	10.8967	0.6204	0.6216	0.7253	82
11.4859	11.0346	10.8967	0.6225	0.6254	0.7253	83
11.8841	12.1412	13.2112	0.6408	0.6429	0.7253	84
13.8164	14.1438	13.2112	0.6598	0.6561	0.7253	85
15.3582	15.4475	16.0016	0.6853	0.6922	0.7253	86
21.0406	20.9006	19.3682	0.6918	0.6942	0.7253	87
21.1426	21.1325	19.3682	0.695	0.697	0.7253	88
22.3399	22.8115	19.3682	0.7201	0.7234	0.7253	89
26.6003	26.3058	19.3682	0.7235	0.7261	0.7253	90

30.3842	30.8722	23.4328	0.7308	0.7279	0.9797	91
31.4402	31.5536	28.3433	0.7685	0.7604	0.9797	92
31.5206	31.5954	28.3433	0.7723	0.7785	0.9797	93
38.3006	36.8453	34.2796	0.7838	0.7797	0.9797	94
41.5408	41.9709	34.2796	0.7913	0.7887	0.9797	95
45.8686	46.0713	41.4604	0.7922	0.7901	0.9797	96
46.4105	46.4655	41.4604	0.8031	0.7999	0.9797	97
62.0402	63.8553	50.1517	0.8109	0.8076	0.9797	98
73.2126	69.4258	73.4311	0.8166	0.8149	0.9797	99
86.1748	92.9336	88.893	0.83	0.8271	0.9797	100

We notice from Figures (3 and 4) that the risk function increases with the passage of time, that is, when time increases. Likewise, for Table (5), the risk to which the patient is exposed also increases with the length of stay in the hospital, meaning that the risk function is directly proportional to time.

10 . Reference

1. Balavarca Y. "**Survival Analysis**", International PhD Program Applied Statistics and Empirical Methods. 8. Biedermann, S., Dette, H., & Zhu, W. (2006). Optimal Designs for Dose–Response Models with Restricted Design Spaces. Journal of the American Statistical Association, 101(474), 747-759
2. Eleuteri,A., Aung,M.S., Taktak,A.F., Damato,B.,& Lisboa,P.J.(2007). Continuous and Discrete Time Survival Analysis: Neural Network Approaches. Annual International Conference of the
3. Gensheimer,M.F.,& Narasimhan,B.(2019). A scalable discrete-time survival model for neural networks. PeerJ 7:e6257. doi.org/10.7717/peerj.6257.
4. Hardin, J. W., & Hilbe, J. M., (2018), “Generalized linear models and extensions”, 4th Edition, Stata press.

5. Ibrahim, J. (2005). Survival Analysis: Introduction. 25. Ismail, A. A., & Sarhan, A. M. (2009). Optimal Design of Step-Stress Life Test with Progressively Type-II Censored Exponential Data. In International Mathematical Forum (Vol. 4, No. 40, pp. 1963-1976).
6. Lamia & Entsar " Structural reliability analysis using the third-moment and Downhill simplex technique " Int. J. Nonlinear Anal. Appl. 13 (2022) 1, 2351-2362 .http://dx.doi.org/10.22075/ijnaa.2022.5933
7. Lamiaa &, Entsar "Analysis of the structural reliability of dental filling under the effect of mouthwash G.U.M . Int. J. Nonlinear Anal. Appl. 12 (2021) No. 2, 2555-2569 http://dx.doi.org/10.22075/ijnaa.2021.5410
8. Madsen, H., & Thyregod, P., (2010), "Introduction to general and generalized linear models", CRC Press, USA.
9. McCullagh, P., & Nelder, J. A., (1989), "Generalized linear models", 2nd edition, Chapman & hall, New York, USA.
10. Miller Jr, R. G. (1980). Survival Analysis Vol. 66). John Wiley & Sons.
11. Mostafa & Entsar " Estimation of the Parameters of the Wiener Degradation Process for the Reliability Analysis Using the Amoebaalgorithm " Mathematical Statistician and Engineering Applications ISS N: 2326-9865 . Article Info Page Number: 1449 – 1460 Publication Issue: Vol 71 No. 3 (2022)
12. Nelder, J.A., & Wedderburn, R.W.M., (1972), "Generalized linear models". J. R. Statist. Soc., A, 135, 370-84
13. Noor & Entsar (2023) " Machine learning methods based on mammogram images to estimate survival times for breast cancer patients" Periodicals of Engineering and Natural Sciences Vol. 11, No. 4, August 2023, pp.133-143
14. Sainani, K. L. (2016). Introduction to Survival Analysis. PM&R, 8(6), 580-585.
15. Tarad , Ala ,a " Parametric and Semiparametric Regression Models (Comparative Study) A thesis Master in Statistics college of Administration &Economics \the University of Mustansiriya
16. Wen ,C, & Chen ,Y.(2020). Discrete-time survival data with longitudinal covariates. John Wiley & Sons, Ltd. Statistics in Medicine Journal;pp.1–14. DOI: 10.1002/sim.8729.
17. Yan, X., & Su, X., (2009), "Linear regression analysis: theory and computing", World Scientific, USA