CHAPTER 3

NON-LINEAR REGRESSION MODELS FOR CENSORED DATA

3.1 INTRODUCTION

Regression modeling of the relationship between an outcome variable and one or more predictor variables is commonly employed in all fields. The popularity of this approach is due to the fact that plausible models may be easily fit, evaluated and interpreted. In this Chapter, we will introduce the parametric Proportional Hazards (PH) model. We will present the Accelerated Failure Time (AFT) model and more detailed discussions of Weibull, log-logistic, log-normal and gamma AFT models. We will apply Cox proportional hazards, Cox proportional hazards with time dependent, Frailty and AFT models, to the Heart attack data. We also give all the corresponding results and compare the models.

3.2 PARAMETRIC PROPORTIONAL HAZARDS MODEL

The parametric proportional hazards model is the parametric versions of the Cox proportional hazards model. The hazard function at time t for the particular patient with a set of p covariates \((x_1, x_2, \ldots, x_p)\), is given as follows:

\[
h(t|x) = h_0(t) \exp(\beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \cdots + \beta_k x_k) = h_0(t) \exp(\beta'x)
\]

The difference between the Cox PH model and Parametric PH models is that the baseline hazard function is assumed to follow a specific distribution when a fully parametric PH model is fitted to the data, whereas the Cox model has no such constraint. The
coefficients are estimated by partial likelihood in Cox model but maximum likelihood in parametric PH model. Other than this, the two types of models are equivalent. Hazard ratios have the same interpretation and proportionality of hazards is still assumed. A number of different parametric PH models may be derived by choosing different hazard functions. The commonly applied models are exponential, Weibull, or Gompertz models.

3.2.1 Weibull PH model

Suppose that survival times are assumed to have a Weibull distribution with scale parameter $\lambda$ and shape parameter $\gamma$, so the survival and hazard function of a $W(\lambda, \gamma)$ distribution are given by

$$
S(t) = e^{-\lambda(t)^\gamma}, 
$$

$$
h(t) = \lambda \gamma \lambda(t)^{\gamma - 1}
$$

with $\lambda, \gamma > 0$. The hazard rate increases when $\gamma > 1$ and decreases when $\gamma < 1$ as time goes on. When $\gamma = 1$, the hazard rate remains constant, which is the special exponential case.

Under the Weibull PH model, the hazard function of a particular patient with covariates $(x_1, x_2, ..., x_p)$ is given by

$$
h(t|x) = \lambda \gamma \lambda(t)^{\gamma - 1} \exp(\beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + ... + \beta_p x_p) = \lambda \gamma \lambda(t)^{\gamma - 1} \exp(\beta'x)
$$

We can see that the survival time of this patient has the Weibull distribution with scale parameter $\lambda \exp(\beta'x)$ and shape parameter $\gamma$. Therefore the Weibull family with fixed $\gamma$ possesses PH property. This shows that the effects of the explanatory variables in the model alter the scale parameter of the distribution, while the shape parameter remains constant.
The corresponding survival function is given by

$$S(t|x) = \exp\{-\exp(\beta'x) \lambda t^r\}$$

(3.1)

After a transformation of the survival function for a Weibull distribution, we can obtain

$$\log\{-\log S(t)\} = \log \lambda + \gamma \log t$$

Then $$\log\{-\log S(t)\}$$ versus $$\log (t)$$ should give approximately a straight line if the Weibull distribution assumption is reasonable. The intercept and slope of the line will be rough estimate of $$\log \lambda$$ and $$\gamma$$ respectively. If the two lines for two groups in this plot are essentially parallel, this means that the proportional hazards model is valid. Furthermore, if the straight line has a slope nearly one, the simpler exponential distribution is reasonable. In the other way, for a exponential distribution, there is $$\log S(t) = -\lambda t$$. Thus we can consider the graph of $$\log S(t)$$ versus $$t$$. This should be a line that goes through the origin if exponential distribution is appropriate.

3.2.2 Exponential PH model

The exponential PH model is a special case of the Weibull model when $$\gamma = 1$$. The hazard function under this model is to assume that it is constant over time. The survival and hazard function are written as

$$S(t) = e^{-\lambda t}, \ h(t) = \lambda.$$  

Under the exponential PH model, the hazard function of a particular patient is given by

$$h(t|x) = \lambda \exp(\beta_1x_1 + \beta_2x_2 + \beta_3x_3 \ldots + \beta_p x_p ) = \lambda \exp(\beta'x).$$

The piecewise exponential model (Breslow,1974) is an extension of the exponential PH model. For the piecewise exponential model, the period of follow-up is divided into
p-intervals \([t_j, t_{j+1}]\), \(j = 1, 2, ..., k\). \(t_1 = 0\). Assume that the baseline hazard is constant within each interval but can vary across intervals, so that \(h_c(t) = \exp(w_j) = \lambda_j\) for \(t_j < t \leq t_{j+1}\), i.e., the baseline hazard function is approximated by a step function.

The piecewise exponential model is given by
\[
\lambda_i(t) = \lambda_j \exp(\beta x_i)
\]
where \(\lambda_i\) is the hazard corresponding to individual \(i\) and \(j\) and \(\exp(\beta x_i)\) is the relative risk for an individual with covariate value \(x_i\) compared to the baseline at any given time.

In the piecewise exponential approach, a log-linear model is used to model both the effects of the covariates and the underlying hazard function. Estimates of the underlying hazard function and the regression parameters can be obtained using maximum likelihood. The maximum likelihood estimates of the baseline hazard function in interval \(i\) for given regression coefficients \(\beta\) is given by
\[
\hat{\lambda}_i = \frac{d_j}{\sum_{l=1}^{n} \hat{R}_{ij} \exp(\beta' x_i) t_l}
\]
where \(d_j\) is the number of events in interval \(j\), \(R_j\) is the risk set entering interval \(j\), and \(t_i\) is the observed survival time for individual \(i\) in interval \(j\). This approach was first studied by (Holford, 1976), and is also the subject of work by (Holford, 1980; Laird and Olivier, 1981).

One of the greatest challenge related to the use of the piecewise exponential model is to find an adequate grid of time-points needed in its construction. One of the advantages of this method is the ability to incorporate time-dependent covariates. If there were any time-dependent covariates, their values at the beginning of each interval could be assigned to the records for that time interval.
3.2.3 Gompertz PH model

The survival and hazard function of the Gompertz distribution are given by

\[ S(t) = e^{-\frac{t}{\theta^b}(e^{\theta t} - 1)}, \quad h(t) = \lambda e^{\theta t} \]

for \(0 \leq t < \infty\) and \(\lambda > 0\). The parameter \(\theta\) determines the shape of the hazard function. When \(\theta = 0\), the survival time then have an exponential distribution, i.e., the exponential distribution is also a special case of the Gompertz distribution. Like the Weibull hazard function, the Gompertz hazard increases or decreases monotonically. For the Gompertz distribution, \(\log h(t)\) is linear with \(t\).

Under the Gompertz PH model, the hazard function of a particular patient is given by

\[ h(t|x) = \lambda \exp(\theta) \exp(\beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 \cdots + \beta_p x_p) = \lambda \exp(\theta t) \exp(\beta'x) \]

It is straightforward to see that the Gompertz distribution has the PH property. But the Gompertz PH model is rarely used in practice.

3.3 ACCELERATED FAILURE TIME MODEL

3.3.1 Introduction

Although parametric PH models are very applicable to analyze survival data, there are relatively few probability distributions for the survival time that can be used with these models. In these situations, the accelerated failure time model (AFT) is an alternative to PH model for the analysis of time to event data. Under AFT models we measure the direct effect of the explanatory variables on the survival time instead of hazard, as we do in the
PH model. This characteristic allows for an easier interpretation of the results because the parameters measure the effect of the correspondent covariate on the mean survival time. Currently the AFT model is not commonly used for the analysis of clinical trial data, although it is fairly common in the field of manufacturing. Similar to the PH model, the AFT model describes the relationship between survival probabilities and a set of covariates.

Under an accelerated failure time model, the covariate effects are assumed to be constant and multiplicative on the time scale, that is, the covariate impacts on survival by a constant factor (acceleration factor).

According to the relationship of survival function and hazard function, the hazard function for an individual with covariates \(X_1, X_2, \ldots, X_p\) is given by

\[
h(t|\mathbf{x}) = \left[1/\eta(\mathbf{x}) \right] h_0\left[t/\eta(\mathbf{x}) \right]
\]

(3.2)

The corresponding log-linear form of the AFT model with respect to time is given by

\[
l = T = \mu + \alpha_1 X_{1t} + \alpha_2 X_{2t} + \cdots + \alpha_p X_{pt} + \sigma \varepsilon_t,
\]

where \(\mu\) is intercept, \(\sigma\) is scale parameter and \(\varepsilon_t\) is a random variable, assumed to have a particular distribution. This form of the model is adopted by most software package for the AFT model.

For each distribution of \(\varepsilon_t\), there is a corresponding distribution for \(T\). The members of the AFT model class include the exponential AFT model, Weibull AFT model, log-logistic AFT model, log-normal AFT model, and gamma AFT model. The AFT models are discussed in details in text books (Collett, 2000; Cox, 1984; Lawless,
The AFT models are named for the distribution of \( T \) rather than the distribution of \( \varepsilon_t \) or \( u_t \).

**Table 3.1: Summary of parametric AFT models**

<table>
<thead>
<tr>
<th>Distribution of ( \varepsilon )</th>
<th>Distribution of ( T )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Extreme Values (1 parameter)</td>
<td>Exponential</td>
</tr>
<tr>
<td>Extreme Values (2 parameters)</td>
<td>Weibull</td>
</tr>
<tr>
<td>Logistic</td>
<td>Log-logistic</td>
</tr>
<tr>
<td>Normal</td>
<td>Log-normal</td>
</tr>
<tr>
<td>Log-Gamma</td>
<td>Gamma</td>
</tr>
</tbody>
</table>

Figure 3.1: Summary of Parametric models

The survival function of \( T_t \) can be expressed by the survival function of \( \varepsilon_t \):

\[
S_t(t) = P(T_t \geq t) = S_{\varepsilon_t} \left( \frac{t - \mu - \alpha}{\sigma} \right).
\] (3.3)

The distribution of \( \varepsilon_t \) and the corresponding distributions of \( T_t \) are summarized in Table (3.1). The effect size for the AFT model is the time ratio. The time ratio comparing two level of covariate \( x_t (x_t = 1) \) \( x_t = 0 \), after controlling all the other covariates.
is $\exp(\mu_1)$, which is interpreted as the estimated ratio of the expected survival times for
two groups. A time ratio above 1 for the covariate implies that this covariate prolongs the
time to event, while a time ratio below 1 indicates that an earlier event is more likely.
Therefore, the AFT models can be interpreted in terms of the speed of progression of a
disease. The effect of the covariates in an accelerated failure time model is to change the
scale, and not the location of a baseline distribution of survival times.

3.3.2 Estimation of AFT model

AFT models are fitted using the maximum likelihood method. The likelihood of
the $n$ observed survival times, $t_1, t_2, \ldots, t_n$ is given by

$$L(u, \mu, \sigma) = \prod_{i=1}^{n} \left\{ f_i(t_i) \right\}^{\delta_i} \left\{ S_i(t_i) \right\}^{1-\delta_i},$$

where $f_i(t_i)$ and $S_i(t_i)$ are the density and survival functions for the $i$th individual at $t_i$
and $\delta_i$ is the event indicator for the $i$th observation. Using equation (3.3), the log-likelihood
function is then given by

$$l(u, \mu, \sigma) = \sum_{i=1}^{n} \left\{ -\delta_i \log(\sigma t_i + \delta_i \log f_i(z_i) + (1 - \delta_i) \log S_i(z_i)) \right\},$$

where $z_i = (t_i - \mu - u_1 x_{1i} - u_2 x_{2i} - \cdots - u_p x_{pi})/\sigma$. The maximum likelihood estimates of the $p + 2$ unknown parameters, $\mu, \sigma, u_1, u_2, \ldots, u_p$, are found by maximizing
this function using the Newton-Raphson procedure, which is the same method used to
maximize the partial likelihood in the Cox regression model.

Several other approaches have been proposed for the estimation and inference on
the AFT model in the literature. Classical semi-parametric approaches to the AFT model
the emphasize estimation of the regression parameters include the method of Buckley and James (1979)and linear-rank-test based estimator (Kalbfleisch and Prentice, 2002). Despite theoretical advances, all these approaches are numerically complicated and difficult to implement, especially when the number of covariates is large.

3.3.3 Weibull AFT model

Suppose the survival time \( T \) has \( \mathcal{W}(\lambda, \gamma) \) distribution with scale parameter \( \lambda \) and shape parameter \( \gamma \). From equation (3.9), under AFT model, the hazard function for the \( i \)th individual is

\[
h_i(t) = \left[ 1/\eta_t(x) \right] h_\varepsilon [t/\eta_t(x)] = 1/[\eta_t(x)]^\gamma \lambda (t)^{\gamma - 1},
\]

where \( \eta_t = \exp (\alpha_1 x_{1i} + \alpha_2 x_{2i} + \cdots + \alpha_p x_{pi}) \) for individual \( i \) with \( p \) explanatory variables. So the survival time for the \( i \)th patient is \( \mathcal{W}(1/[\eta_t(x)]^\gamma \lambda, \gamma) \). The Weibull distribution has the AFT property.

If \( T_i \) has a Weibull distribution, then \( \varepsilon_i \) has an extreme value distribution (Gumbel distribution). The survival function of Gumbel distribution is given by

\[
S_{\varepsilon_i}(\varepsilon) = \exp(-\exp(\varepsilon)).
\]

From equation (3.3), the AFT representation of the survival function of the Weibull model is given by

\[
S_i(t) = \exp \left[ - \exp \left( \frac{-\mu - \alpha_1 x_{1i} - \alpha_2 x_{2i} - \cdots - \alpha_p x_{pi}}{\sigma} \right)^{1/\gamma} \right]. \tag{3.4}
\]

From equation (3.1), the PH representation of the survival function of the Weibull model is given by
\[ S_{t}(t) = e^{- \left( e^{(\beta_{1}x_{1t} + \beta_{2}x_{2t} + \cdots + \beta_{p}x_{pt})t} \right)} \]  

(3.5)

Comparing the above two formulas (3.4) and (3.5), we can easily see that the parameter \( \lambda, \gamma, \beta_j \) in the PH model can be expressed by the parameters \( \mu, \gamma, \alpha_j \) in the AFT model:

\[ \lambda = \exp(-\mu/\sigma), \quad \gamma = 1/\sigma, \quad \beta_j = -\alpha_j/\sigma. \]  

(3.6)

The AFT representation of hazard function of the Weibull model is given by

\[ h_{t}(t) = \frac{1}{\sigma t^{\gamma-1}}e^{-\left(\frac{\mu - \alpha_{1}x_{1t} - \alpha_{2}x_{2t} - \cdots - \alpha_{p}x_{pt}}{\sigma}\right)}. \]  

(3.7)

3.3.4 Log-logistic AFT model

One limitation of the Weibull hazard function is that it is a monotonic function of time. However, the hazard function can change direction in some situations. The log-logistic survival and hazard function are given by

\[ S(t) = \frac{1}{1 + e^{\theta t^k}}, \quad h(t) = \frac{e^{\theta k} k^{-1}}{1 + e^{\theta t^k}} \]

where \( \theta \) and \( k \) are unknown parameters and \( k > 0 \). When \( k \leq 1 \), the hazard rate decreases monotonically and when \( k > 1 \), it increases from zero to a maximum and then decreases to zero.

Suppose that the survival times have a log-logistic distribution with parameter \( \theta \) and \( k \), then from equation (3.2), under the AFT model, the hazard function for the \( i \)th individual is

\[ h_{i}(t) = \left[1/\eta_{i}(x)\right]h_{x}[t/\eta_{i}(x)] = \frac{e^{\theta - k} \eta_{i}k t^{k-1}}{1 + e^{\theta - k} \eta_{i}t^{k}} \]
where \( \eta_i = \exp \left( \alpha x_{1i} + \alpha x_{2i} + \cdots + \alpha x_{pi} \right) \) for individual \( i \) with \( p \) explanatory variables. So the survival time for the \( i \)th individual has a log-logistic distribution with parameter \( \theta - k \) \( \eta_i \) and \( k \), log-logistic distribution has AFT property.

If the baseline survival function is \( S_0(t) = \left\{1 + e^{\theta t^k}\right\}^{-1} \), where \( \theta \) and \( k \) are unknown parameters, the baseline odds of surviving beyond time \( t \) are given by

\[
\frac{S_0(t)}{1 - S_0(t)} = e^{-\theta t^k}.
\]

The survival time for the \( i \)th individual also has a log-logistic distribution, which is

\[
S_i(t) = \frac{1}{1 + e^{\theta t^k - k}} \quad \text{(3.8)}
\]

Therefore, the odds of the \( i \)th individual surviving beyond time \( t \) is given by

\[
\frac{S_i(t)}{1 - S_i(t)} = e^{\mu_i - \theta t^k} \quad \text{(3.9)}
\]

We can see that the log-logistic distribution has the proportional odds (PO) property. So this model is also a proportional odds model, in which the odds of an individual surviving beyond time \( t \) as expressed as

\[
\frac{S_i(t)}{1 - S_i(t)} = \exp \left( \beta_1 x_{1i} + \beta_2 x_{2i} + \cdots + \beta_p x_{pi} \right) \frac{S_0(t)}{1 - S_0(t)}.
\]

In any two group study, using (3.9), the log (odds) of the \( i \)th individual surviving beyond time \( t \) is given by

\[
\log \left[ \frac{S_i(t)}{1 - S_i(t)} \right] = \beta x_{1i} - \theta - k
\]
where \( x_i \) is the value of a categorical variable which takes the value one in one group and zero in the other group. A plot of \( \log \left[ \frac{(1 - S(t))/S(t)}{\mu} \right] \) versus \( t \) should be linear if log-logistic distribution is appropriate. Therefore we can check the suitability of log-logistic distribution using the PO property.

If \( T_i \) has a log-logistic distribution, then \( \varepsilon_i \) has a logistic distribution. The survival function of logistic distribution is given by

\[
S_{\varepsilon_i}(\varepsilon) = \frac{1}{1 + \exp(\varepsilon)}.
\]

Using equation (3.3), the AFT representation of survival function of the log-logistic model is given by

\[
S_i(t) = \left[ 1 + \frac{t}{\theta} e^{\theta} \right]^{-1} \left[ \frac{\mu - \mu_1 x_{i1} - \mu_2 x_{i2} - \cdots - \mu_p x_{ip}}{\sigma} \right]^{-1}. \tag{3.10}
\]

Comparing the formula (3.8) and (3.10), we can easily find \( \theta = -\mu/\sigma, \ k = \sigma^{-1} \).

According to the relationship of survival and hazard function, the hazard function for the \( i \)th individual is given by

\[
h_i(t) = \frac{1}{\sigma} \left[ 1 + \frac{t}{\theta} e^{\theta} \right]^{-1} \left[ \frac{\mu + \mu_1 x_{i1} + \mu_2 x_{i2} + \cdots + \mu_p x_{ip}}{\sigma} \right]^{-1}. \tag{3.11}
\]

3.3.5 Log-normal AFT model

If the survival time times are assumed to have a log-normal distribution, the baseline survival function and hazard function are given by

\[
S_0(t) = 1 - \Phi \left( \frac{t_0 - \mu}{\sigma} \right), \quad h_0(t) = \frac{\phi \left( \frac{t_0 - \mu}{\sigma} \right)}{1 - \Phi \left( \frac{t_0 - \mu}{\sigma} \right)} \sigma,
\]

68
where $\mu$ and $\sigma$ are parameters, $\phi(x)$ is the probability density function and $\Phi(x)$ is the cumulative density function of the standard normal distribution. The survival function for the $i$th individual is

$$S(t) = S_0(t/\eta_i) = 1 - \Phi\left(\frac{\log t - \alpha' x_i - \mu}{\sigma}\right),$$

where $\eta_i = \exp\left(\alpha_1 x_{i1} + \alpha_2 x_{i2} + \cdots + \alpha_p x_{ip}\right)$. Therefore the log survival time for the $i$th individual has normal $(\mu + \alpha' x_i, \sigma)$. The log-normal distribution has the AFT property.

$$\Phi^{-1}[1 - S(t)] = 1/\sigma (\log t - \alpha' x_i - \mu),$$

where $x_i$ is the value of a categorical variable which takes the value one in one group and zero in the other group. This implies that a plot of $\Phi^{-1}[1 - S(t)]$ versus $u$ will be linear if the log-normal distribution is appropriate.

### 3.3.6 Gamma AFT model

There are two different gamma models discussed in survival analysis literature. The standard (2-parameter) and the generalized (3-parameter) gamma model. The gamma model means the generalized gamma model. The probability density function of the generalized gamma distribution with three parameters $\lambda, \alpha, \gamma$ is defined by

$$f(t) = \frac{\alpha \lambda^\alpha}{\Gamma(\gamma)} t^{\alpha-1} \exp\left\{-\left(\frac{\lambda}{\gamma}\right)^\alpha \right\} t > 0, \gamma > 0, \lambda > 0, \alpha > 0,$$

where $\gamma$ is the shape parameter of the distribution. The survival function and the hazard function do not have a closed form for the generalized gamma distribution. The exponential, Weibull and log-normal models are all special cases of the generalized gamma model. It is easily to seen that this generalized gamma distribution becomes the
exponential distribution if $\alpha = \gamma = 1$, the Weibull distribution if $\gamma = 1$, and the log-normal distribution if $\gamma \to \infty$. The generalized gamma model can take on a wide variety of shapes except for any of the special cases.

3.4 SEMI PARAMETRIC MODEL (COX REGRESSION MODEL)

For the diagnosis of a disease, medical doctors investigate the cause or the other characteristics of a disease. For example, is the heart patient has the disease of high blood pressure? Or family history of diabetic related to the development of diabetic disease? In this case, high blood pressure and family history are referred to as covariates or risk factors or explanatory variables. Now-a-days, the identification of the most important risk factors is becoming the important task for handling the disease. Regression analysis is generally used for identifying the risk factors. But due to the presence of censoring in time to event data, ordinary regression models are not used on time to event data. For this purpose, in survival analysis, Cox regression model/ Cox proportional hazard (PH) model (Cox, 1972) is widely used. The proportional hazards regression model is very popular due to the easy concept and accessibility software (Altman and Stavola, 1994; Lin, 1991; Aalen, 1989; Lagakos and Schoenfeld, 1984; Solomon, 1984; Bryson, Jonhson, 1981). Like logrank test, it is also based on the proportional hazards assumption.

If a set of covariates is represented by $X_t = (x_{1t}, x_{2t}, \ldots, x_{kt})$, famous Cox regression model is

$$h_t(t, x_t) = h_0(t) \exp \left( \beta x_t \right)$$

$$h_t(t, x_t) = h_0(t) \exp \left( \sum_{i=1}^{k} \beta_i x_{it} \right), \quad j = 1, 2, \ldots, k$$

70
where $h_0(t)$ is an unspecified baseline hazard function and is a function of time only, assume to be same for all subjects, $\beta_j$ are unknown parameters, describe the importance of covariates and $x_{ij}$ are the values measured on subject $i$ at time zero. The baseline hazard describes the shape of the distribution while $\exp(\beta_1 x_i)$ gives the level of each individual’s hazard. The model based on the assumption that independent covariates effect the hazard in a multiplicative way (Marubini, Valsecchi, 1995). $x$ is a vector of covariates of interest. $x$ may be discrete (sex, marital status), continuous (blood pressure) or the mixture of discrete and continuous factors (interaction height and sex). The main advantage of Cox PH model is that we can estimate the parameter $\beta$, without having to estimate $h_0(t)$. This fact makes the model as a semi-parametric.

The model can also be expressed as

$$y_i = \sum_{j=1}^{k} \beta_j x_{ij}$$

If we take $y_i = l_i \left[ \frac{h_1(t, x_i)}{h_0(t)} \right]$, the above equation becomes the multiple regression equation.

$$y_i = \sum_{j=1}^{k} \beta_j x_{ij}$$

It is declared that the hazards for two individuals ($x_1$ and $x_2$) are assumed to be proportional

$$H = \frac{h_1(t)}{h_2(t)} = \exp(\beta_1 (x_1 - x_2))$$

Cox assumes that hazard ratio comparing two specifications of covariates is constant over time. Example of such covariates are sex of the patient, treatment etc. Consider a single
covariate $x$, assuming two values 0 and 1. Where $x = 0$ if the patient receives treatment 1 and $x = 1$, if the patient receives treatment 2. In terms of proportional hazard, this can be written as $h_2(t) = e^\beta h_1(t)$. If $e^\beta < 1$, the effect of treatment 2 is more than treatment 1. Both have the same effects if $e^\beta = 1$. If $e^\beta > 1$, treatment 1 is superior to treatment 2. $\beta$ is the log hazard ratio.

The assumption of proportional hazard in Cox regression means that the base line hazard function is only a function of time (t), where the exponential part involves $x_1$'s but independent of t. If it is not possible then $x_1$'s called the time dependent covariates otherwise time independent covariates. Patient performance during the treatment period is an example of time dependent covariates. Several procedures have been suggested for checking this assumption (Petitt, 1989; Kay, 1984; Andersen, 1982). An easy and simple method of checking the proportional hazards assumption in the presence of covariates is the plotting method. Suppose $x_1$ and $x_2$ are two different set of covariates, draw the plot of log $[-\log S(t, x_1)]$ and log $[-\log S(t, x_2)]$ against time (function of time), where $S(t, x)$ is the survival function. If the two curves are parallel, there is no reason to believe that the proportional hazards assumption is not satisfied. The method is easy for few covariates but for large number it becomes tedious. Except the plot of cumulative hazard functions for checking the assumption of proportional hazard model, there is another method which is based on the concept of introducing the time dependent variable $t\circ l(t)$ as the independent variable in to the model and centered about a point $t^*$.

$$h(t) = h_0(t) e^{\beta_1 x + \beta_2 (l(t) - l(t^*))}$$ (3.16)
Test $\beta_2 = 0$, if it is true, proportional hazard assumption is satisfied. If it is not, time dependent can be used.

Since $h_c(t)$ is free from parametric assumption, it is not possible to apply the full likelihood function for estimating $\beta$'s. For the model, Cox [1972, 1975] suggested an estimation procedure, in which the analysis concentrates only on the effect of covariates and leaving $h_c(t)$ completely unspecified. $\beta$ coefficients in Cox proportional hazards model can be estimated by using the partial likelihood method, which is a function of observed survival times and unknown parameters.

$$I_k(\beta) = \prod_{t=1}^{n} \frac{\varepsilon_{j=1}^{k} \beta_j x_{ij}}{\sum_{t \in H(t)} \varepsilon_{j=1}^{k} \beta_j x_{ij}}$$

where $t_{(1)}, t = 1, 2, ..., n$ denote the ordered exact failure times and $H(t_{(1)})$ consists of all individuals whose survival times are at least $t_{(1)}$. The summation in the denominator of likelihood is the sum of the values of $\varepsilon_{j=1}^{k} \beta_j x_{ij}$ over all individuals who are at risk at this time. The likelihood depends only on the rank of failure times and the variance of the partial likelihood is larger than the variance of complete likelihood.

Tied survival times can be handled by using the likelihood form of (Breslow, 1974). This is simple to understand and is more suitable in a situation when the number of tied observations at any death time is not too large.

$$I_k(\beta) = \prod_{t=1}^{n} \left( \frac{\varepsilon_{j=1}^{k} \beta_j x_{ij}}{\sum_{t \in H(t)} \varepsilon_{j=1}^{k} \beta_j x_{ij}} \right)^{\xi_t}$$

where $\xi_t$ denotes the number of tied survival times. In 1981, Tsiatis investigated its properties (Tsiatis, 1981). The other model dealing with tied situation are (Efron, 1977; Cox, 1972;
Kalbfleisch and Prentice, 1980). In case of large tied survival times, Cox’s assumes the discrete survival times and generalized the model by introducing the concept of logistic transformation

\[
\frac{h(t)d}{1 - h(t)d} = \frac{h_0(t)d}{1 - h_0(t)d} e^{\sum_{j=1}^{k} \beta_j x_{ij}}
\]  

(3.19)

All these procedures are used for Cox Proportional hazards model. If its assumption failed, one cannot use the procedure. There are several reasons for the failure of assumption out of which the most common reason is the involvement of time dependent covariates. Example of time dependent covariates are, status of cigarette smoking may change during the study period, the cholesterol level of a patient changes during the study period, regular examination of patients in a clinic is also an example of a time dependent covariate. In time dependent Cox regression, the values of variable changes over the period or time. Time dependent covariates are also referred to as updated covariates (Altman and Stavola, 1994) and are of two types, external and internal variables.

External variable is a variable whose values in most cases are known in advance. Such variables do not require the survival of patients. Simple examples of external variable are the age of a patient, dose of a drug. The measurement of internal variable is only possible, if the patient is alive. Examples are systolic blood pressure, white blood cell count etc. For the analysis of time dependent covariates, extended Cox model may be used, called time dependent or non-proportional hazard Cox model.

Consider a set of q covariates, out of which \( q_1 \) are time independent and \( q_2 \) are time dependent covariates. The modified form of the Cox proportional hazards model is obtained by dividing the exponential part in to time independent and time dependent parts.
Like the Cox proportional hazard model, the regression coefficients for the time dependent model is also obtained through the maximum likelihood procedure. The likelihood for time dependent is the same as that for time-independent except that $z$ is replaced to $x(t)$, so

$$l_{i_k}(\beta) = \prod_{i=1}^{n} \left[ \frac{\varepsilon_i \left( \sum_{j=1}^{k} \beta_j x_{ij}(t_i) \right)}{\sum_{l \in i \in H(t_i)} \varepsilon_l \left( \sum_{j=1}^{k} \beta_j x_{il}(t_i) \right)} \right]^{I_i}$$ (3.21)

3.5 FRAILTY MODELS

The concept of frailty provides a suitable way to introduce random effects in the model to account for association and unobserved heterogeneity. In its simplest form, a frailty is an unobserved random factor that modifies multiplicatively the hazard function of an individual or a group or cluster of individuals.

3.5.1 Introduction

Vaulpe et al. (1979) introduced the term frailty and used it in univariate survival models. Clayton (1978) promoted the model by its application to multivariate situation on chronic disease incidence in families. A random effect model takes into account the effects of unobserved or unobservable heterogeneity, caused by different sources. The random effect, called frailty and denoted here by $Z$ is the term that describes the common risk or the individual heterogeneity, acting as a factor on the hazard function. Two categories of frailty models can be pointed out. The first one is the class of univariate frailty models that consider univariate survival times. The second one is the class of multivariate frailty models that take into account multivariate survival times.
3.5.2 Univariate frailty models

Univariate frailty models take into account that the population is not homogeneous. Heterogeneity may be explained by covariates, but when important covariates have not been observed, this lead to unobserved heterogeneity. Vaupel et al. (1979) introduced univariate frailty models (with a gamma distribution) into survival analysis to account for unobserved heterogeneity or missing covariates in the study population. The idea is to suppose that different patients possess different frailties and patients more “frail” or “prone” tend to have the event earlier that those who are less frail. The model is represented by the following hazard given the frailty:

\[ \lambda(t|Z,X) = Z \lambda(t|X) \]

\( \lambda(t|X) \) can be equal to the baseline hazard function \( \lambda_b(t) \), or when consider covariates \( \lambda(t|X) \) may be equal to \( \lambda_b(t) \exp(\beta^T X) \) (in a Cox regression model). The baseline hazard function \( \lambda_b(t) \) can be chose non-parametrically, or parametrically (Weibull, exponential, Gompertz, …). An important point is that the frailty \( Z \) is an unobservable random variable varying over the sample which increases the individual risk if \( Z > 1 \) or decreases if \( Z < 1 \).

The model can also be represented by its conditional survivor function:

\[ S(t|Z,X) = \exp\left(-Z \int_{t_c}^t \lambda(u|X) \, du \right) = \exp\left(-Z \Lambda(t|X) \right), \]

where \( \Lambda(t|X) = \int_{t_c}^t \lambda(u|X) \, du \). \( S(t|Z,X) \) represents the fraction of individuals surviving until time \( t \) given \( Z \) and given the vector of observable covariates \( X \). Note that until now the model is described at individual level, but this individual model is not observable. That is
the reason why it is essential to consider the model at a population level. The survival of the total population is the mean of the individual survival functions. Many calculations can be done based on the Laplace transform. Hougaard (1984) demonstrated the importance of the Laplace transform of these calculations. The Laplace transform of a random variable \( Z \) is defined as

\[
L(s) = \int \exp(-sz) g(z) dz = \mathbb{E}[\exp(-sz)]
\]

where \( g(z) \) is the density of \( Z \). The integral is over the range of the distribution. The marginal survivor function can be calculated by

\[
S(t|X) = \int S(t|Z,X) g(z) dz = \mathbb{E}[S(t|Z,X)] = L(\Lambda(t|X)).
\]

An important point is the identifiability of univariate frailty models. Univariate frailty models are not identifiable from the survival information alone. However, Elbers and Ridder (1982) proved that a frailty model with finite mean is identifiable with univariate data, when covariates are included in the model. Many distributions can be chosen for the frailty, but the most common frailty distribution is the gamma distribution. The gamma distribution has been widely applied as a mixture distribution (Clayton, 1978; Hougaard, 2000; Oakes, 1982; Vaupel et al., 1979; Yashin et al., 1995). From a computational and analytical point of view the gamma distribution is convenient, because it is easy to derive the closed form expressions of survival, density and the hazard function. This is due to the simplicity of the Laplace transform, which is the reason why this distribution has been used in most applications. In gamma distribution, \( \theta \) is a scale parameter and \( \beta \) is a shape parameter. For identifiability, we suppose \( \theta = \beta \) which implies \( \nu = 1 \) and \( \nu = \frac{1}{\theta} \).
An other distribution which can be chosen for the frailty is the positive stable distribution (Hougaard, 1986). A distribution is strictly stable if the sum of independent random variables from the distribution normalized follows the same distribution. Suppose \( Z_1, Z_2, Z_3, \ldots, Z_n \) i.i.d, the distribution of the sum of \( Z_1, Z_2, Z_3, \ldots, Z_n \) is stable if for each \( n \), there exists a constant \( c_n \), with \( U(Z_1 + Z_2 + Z_3 + \ldots + Z_n) = U(c_nZ_1) \) where \( U(Z) \) means the distribution of \( Z \). The constants satisfy \( c_n = n^{1/\alpha} \), for some \( \alpha \in (0, 2] \). For \( \alpha = 2 \), the stable distribution has finite variance and is the normal distribution. For \( \alpha = 1 \), the degenerate distribution is obtained. The stable distribution on the positive numbers has \( \alpha \in (0, 1] \) and apart from scale factors have Laplace transform:

\[
L(s) = E[\exp(-sZ)] = \exp(-s^{\alpha})
\]

\((s \geq 0)\). This distribution is denoted \( P(\alpha, \alpha, 0) \). Note that the frailty model using this distribution is not identifiable in the univariate case, because the mean does not exist. Unidentifiability is also easily seen from the marginal survival function: \( S(t|X) = \exp((-\Lambda(t) \exp(X))^{\alpha}) = \exp(-\alpha \Lambda(t) \exp(X)) \), where the frailty parameter (\( \alpha \)) acts as a multiplicative factor which is confounded by \( \Lambda(t) \).

Other distributions which are sometimes applied for the frailty distribution are the well-known normal, the lognormal (McGilchrist and Aisbett, 1991), the three parameter distribution (PVF) (Hougaard, 1986), the compound poisson distribution (Aalen, 1988, 1992) and inverse Gaussian distribution. The effect of different frailty distribution is investigated by Congdon (1995).

The role of shared frailty is more useful when we consider multivariate survival times.
3.5.3 Multivariate frailty models

A very common situation in survival analysis is clustered or repeated data. Clustered data are for instance data where individuals are divided in groups like family or study centres. Repeated data are seen in case of longitudinal data. Concerning multiple recurrences of an event for the same individual. The difficulty of working with this kind of data is due to the dependence of individuals within groups, or repeated measures within individuals. The dependence usually arises because individuals in the same group are related to each other or because of the recurrence of an event for the same individual. Multivariate frailty models have been used frequently for modeling dependence in multivariate time-to-event data (Clayton, 1978; Hougaard, 2000; Oakes, 1982; Yashin et al., 1995). The aim of the frailty is to take into account the presence of the correlation between the multivariate survival times.

**Constant shared frailty models**

In this situation, individuals \( j \) in a group \( t \) are supposed to share the same frailty \( Z_t \). The conditional hazard for individual \( j \) in group \( t \) is:

\[
\lambda(t_t | Z_t) = Z_t \lambda(t_t),
\]

where \( \lambda(t_t) = \lambda_u(t_t) \exp(\beta X_t) \) in the cox-regression model. The \( Z_t \) are independent identically distributed following a chosen distribution, like in the univariate frailty models. This model is therefore an extension of the preceding described model. The model assumes that all time observations are independent given the values of the frailties. In other words, it is a conditional independence model. The value of \( Z \) is constant over time and common to the individuals in the group and thus responsible for creating dependence. The interpretation of this model is that the between-groups variability (the random
variation of \( Z \) leads to different risks for the groups, which then show up as dependence within the group. In the case of gamma distribution for \( Z \), remember that \( E = 1 \) and \( \psi = \frac{1}{\theta} \). So, small value of \( \theta \) reflect a greater degree of heterogeneity among groups and a stronger association within groups. The association between group members as measured by kendall’s \( \tau \) is \( \tau = \frac{1}{1+2\theta} \), and large value of \( \theta \) corresponds to the case of independence.

Note that the frailty models with multivariate survival data are identifiable in almost all cases. It is assumed that there is independence between groups and between the times for the same value of \( i \), owing to the common value \( Z \) of \( Z \). Thus if the \( Z \) do not vary then there is independence between the time observations.

**Gamma frailty model**

A first and common approach is to define the hazard function as:

\[
\lambda(t_i | Z_i) = Z_i \lambda_\theta(t_i) \exp(\beta^T X_i), \quad i = 1, 2, \ldots; \quad j = 1, 2, \ldots, k_i
\]

which is the hazard function of the \( j^{th} \) individual group \( i \) given the frailty of group \( i \) (\( Z_i \)), where \( \lambda_\theta(t_i) \) is an arbitrary baseline hazard rate and \( X_i \) is the corresponding covariate vector. The frailty \( Z \) is supposed to follow a gamma distribution. The joint survival function for the \( k_i \) individuals within the \( i^{th} \) group is easily written by:

\[
S\{t_{i1}, \ldots, t_{ik_i}\} = P_{R}(T_{i1} > t_{i1}, \ldots, T_{ik_i} > t_{ik_i})
\]

\[
= \int_{\mathbb{R}} \prod_{j=1}^{k_i} P_{R}(T_i > t_i | Z_i) \varphi(z_i) \, dz_i
\]

(3.22)
\[
= 1 + \frac{1}{\theta} \sum_{j=1}^{K_i} \Lambda_{c}(t_i) e^{- (\beta^T X_i) t_i}
\]

In this model, the estimates of \( \hat{\beta}, \hat{\theta}, \Lambda_{c}(t) \) are obtained by using the EM (Expectation Maximization) algorithm (Dempster et al., 1977). The EM algorithm is the main tool for estimation in frailty models in a frequentist framework and provides a means of maximizing complex likelihoods. The likelihood considered is the full likelihood we would have if the frailties were observed. This likelihood is easily manipulable and written as follows: \( L_f = l_1(\theta) + l_2(\Lambda_{c}) \). In the E step the expected value of the full likelihood is completed given the current estimates of the parameters and the observable data. In the M step the estimates of the parameters which maximize the expected value of the full likelihood form the E step are obtained. For more details see Klein and Moeschberger (1997).

If one assumes a parametric form for \( \Lambda_{c}(t_i) \), then, ML estimates are available by maximizing the log likelihood directly. In this following parametric example, the weibull distribution is chosen. This model is called the gamma-weibull frailty model:

\[
L_i = \text{Pr}((t_{1i},d_{1i}),\ldots,(t_{1i},d_{1i})) = \int \text{Pr}((t_{1i},d_{1i}),\ldots,(t_{1i},d_{1i})|Z_i) g(z_i) dz_i
\]

\[
= \prod_{j=1}^{K_i} \int \sum_{t_i} [z_i \Lambda_{c}(t_i) \exp(\beta^T X_i)]^{d_i} \exp(-z_i \Lambda_{c}(t_i) \exp(\beta^T X_i)) g(z_i) dz_i
\]
Because in the Weibull situation, \( \lambda_0(t_i) = \alpha t_i^{\mu-1} \) and the corresponding cumulative baseline hazard \( \lambda_0(t_i) = \beta t_i^\mu \) the final expression of the likelihood is then easily derived, and also the log likelihood. Usually the loglikelihood is directly maximized using Newton-Raphson procedures and estimates of the variability of the parameter estimates are obtained by inverting the information matrix.
3.6 APPLICATION TO REAL DATA

Non-Linear Regression Models for Heart Attack Data – An empirical comparison

3.6.1 Introduction

Cardiovascular disease is a broad term used to describe a range of diseases that affect heart or blood vessels. The various diseases that fall under the umbrella of cardiovascular disease include coronary artery disease, heart attack, heart failure, high blood pressure and stroke (Hayes, 2007). Heart failure occurs when the pumping action of the heart cannot provide enough blood to the part of the body as it is needed. Congenital heart disease is a problem with the structure of the heart because of a birth defect (Steinberger et al., 2003). The heart works hard to pump blood around the body and causes blood to flow at increased pressure through the blood vessels, causing risk for a heart attack or stroke. High blood pressure is a major risk factor for coronary heart disease and stroke (Gibbs et al., 2000). Stroke occurs when a blood vessel supplying blood to one part of the brain becomes blocked. As a result, no oxygen can get to this part of the brain and leads to death. Symptoms of stroke include weakness in the arm, leg, hand or face, sudden blindness, difficulty in speaking and loss of balance (Schoenstad, 2008)

An estimated 17.5 million people died from cardiovascular disease in 2005, representing 30 percent of all global deaths. Of these deaths, 7.6 million were due to heart attacks and 5.7 million due to stroke and 4.2 million due to hypertension and other heart conditions. About 80 percent of these deaths occurred in low and middle income countries. World Health Organisation, 2008 cautioned that by 2015, an estimated 20 million people will die from cardiovascular disease (Boutayeb, 2006). American Heart Association, (2007) reported that cardiovascular disease is the single largest cause of death among
women in the United States and worldwide, accounting for one-third of all deaths. In fact, more women than men die every year of cardiovascular disease. There are about 4,600 babies born with congenital heart disease each year and 12 million people are currently affected by rheumatic fever and rheumatic heart disease. Two-thirds are children between 5 and 15 years of age (British Heart Foundation, 2007).

Global burden disease study reported that of a total of 9.4 million deaths in India in 1990, cardiovascular disease caused 2.3 million deaths (25 percent), 1.2 million deaths were due to coronary heart disease and 0.5 million due to stroke. It has been predicted that by 2020, there would be a 111 percent increase in cardiovascular disease deaths in India. Almost 2.6 million Indians are predicted to die due to coronary heart disease which constitutes 54.1 per cent of all cardiovascular diseases deaths in India by 2020. Cardiovascular diseases are the major cause of mortality and disease in the Indian subcontinent, causing more than 25 per cent of deaths. It has been predicted that these diseases will increase rapidly in India and this country will be host to more than half the cases of heart diseases in the world within the next 15 years (Gupta et al., 2008). In India heart disease occurs five to ten years earlier with 50 per cent of heart attacks below the age of 55 and 25 per cent below the age 40 (Gupta, 2004).

The aim is to compare the performance of Cox proportional hazard model, Cox time dependent model, Frailty model and AFT using Heart attack data.

3.6.2 Description of the data set

The data obtained from the website, ftp://ftp.wiley.com/public/scitech_med/survival and http://www.umass.edu/statdata/statdata. The data from the Worcester Heart Attack Study (WHAS) have been provided by Goldberg (1989) of the Department of Cardiology
at the University of Massachusetts Medical School. Data have been collected during 13 one-year periods (1975-1988), on all myocardial infarction (MI) patients admitted to hospital in the Worcester, Massachusetts Standard Metropolitan Statistical Area. Event is coded as 1 and censoring is coded as 0.

3.6.3 Description of Covariates

The subsets of covariates used, with their codes and values for provided in the following table:

<table>
<thead>
<tr>
<th>COVARIATES</th>
<th>DESCRIPTION</th>
<th>CODES/UNITS</th>
</tr>
</thead>
<tbody>
<tr>
<td>AGE</td>
<td>Age</td>
<td>years</td>
</tr>
<tr>
<td>SEX</td>
<td>Gender</td>
<td>0 = Male, 1 = Female</td>
</tr>
<tr>
<td>CPK</td>
<td>Peak Cardiac Enzyme</td>
<td>Int. units</td>
</tr>
<tr>
<td>SHO</td>
<td>Cardiogenic shock</td>
<td>0 = No, 1 = Yes</td>
</tr>
<tr>
<td></td>
<td>complications</td>
<td></td>
</tr>
<tr>
<td>CHF</td>
<td>Left Heart Failure</td>
<td>0 = No, 1 = Yes</td>
</tr>
<tr>
<td></td>
<td>Complications</td>
<td></td>
</tr>
<tr>
<td>MIORD</td>
<td>MI Order</td>
<td>0 = First, 1 = Recurrent</td>
</tr>
<tr>
<td>YRGRP</td>
<td>Grouped Cohort Year</td>
<td>1 = 1975 &amp; 1978</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2 = 1981 &amp; 1984</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3 = 1986 &amp; 1988</td>
</tr>
</tbody>
</table>
### 3.6.4 Model Results

Table 3.3 Parameter Estimates under different Models

<table>
<thead>
<tr>
<th>COVARIATES</th>
<th>COX PH</th>
<th>COX TIME DEPENDENT</th>
<th>FRAILTY</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( \beta )</td>
<td>S.E.</td>
<td>( \beta )</td>
</tr>
<tr>
<td>AGE</td>
<td>.035*</td>
<td>.006</td>
<td>.033*</td>
</tr>
<tr>
<td>SEX</td>
<td>.045</td>
<td>.134</td>
<td>.008</td>
</tr>
<tr>
<td>SHO</td>
<td>1.860*</td>
<td>.217</td>
<td>1.906*</td>
</tr>
<tr>
<td>CHF</td>
<td>.563*</td>
<td>.143</td>
<td>.564*</td>
</tr>
<tr>
<td>MIORD</td>
<td>.269*</td>
<td>.132</td>
<td>.211</td>
</tr>
<tr>
<td>YRGRP</td>
<td>-.205*</td>
<td>.094</td>
<td>-.150</td>
</tr>
</tbody>
</table>

| DEVIANACE  | 2662.699 | 2659.617 | 2662.277 |

*P < 0.05 Significant
Table 3.4 Different approaches of Accelerated Failure Time models

<table>
<thead>
<tr>
<th>VARIABLES</th>
<th>WEIBULL</th>
<th>EXPONENTIAL</th>
<th>LOG-LOGISTIC</th>
<th>LOG-NORMAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>AGE</td>
<td>-0.157*</td>
<td>-0.078*</td>
<td>-0.157*</td>
<td>-0.169*</td>
</tr>
<tr>
<td>SEX</td>
<td>-0.829</td>
<td>0.266</td>
<td>-1.199</td>
<td>-1.417</td>
</tr>
<tr>
<td>SHO</td>
<td>-10.087*</td>
<td>-5.388*</td>
<td>-9.646*</td>
<td>-9.819*</td>
</tr>
<tr>
<td>CHF</td>
<td>-2.323*</td>
<td>-0.570*</td>
<td>-2.372*</td>
<td>-2.361*</td>
</tr>
<tr>
<td>MIORD</td>
<td>-0.166</td>
<td>0.409</td>
<td>-0.876</td>
<td>-1.247</td>
</tr>
<tr>
<td>YRGRP</td>
<td>0.707</td>
<td>0.372*</td>
<td>0.419</td>
<td>0.502</td>
</tr>
<tr>
<td>DEVIANCE</td>
<td>727.616</td>
<td>1068.560</td>
<td>730.611</td>
<td>727.875</td>
</tr>
</tbody>
</table>

3.6.5 Results

The non-linear regression models were fitted using STATA 12 and the results are presented in Table 3.3 and Table 3.4. From the Table 3.3, we see that the three covariates namely AGE, SHO, CHF are significantly associated with the survival time under all the model. Among the models, Time dependent Cox model has the lowest level of deviance compared to all other models. It is further noted that all other models have significantly higher deviance compared to Time dependent Cox model.
3.6.6 Summary

From the Table 3.3 it is found that the, SEX is not significant to the survivorship. Controlling for other covariates the risk of death increases by 3.6% \[e^{\lambda(0.035)} - 1\] as a patient’s age increases by 1 year. Controlling for other covariates, the risk of death of patients with recurrent heart attack is 1.3 \[e^{\lambda(0.269)}\] times the risk of patients with first heart attack. Using Cox PH model for all covariates the result showed significant p-values except the covariate SEX. When Extended Cox PH model was used for the time dependent covariate, Peak Cardiac Enzyme (CPK). The result is based on CPK as a time dependent covariates, after adjusting all the covariates. We observed that the result of time dependent covariates is better than the results of Cox PH Model. The application of frailty model to the same data assuming individual heterogeneity is also considered. For Frailty model, we see that all the covariates are statistically significant except SEX and MIORD. This leads to the conclusion that the individual heterogeneity in SEX and MIORD differs significantly between the patients. Among the models, Time dependent Cox model has the lowest level of deviance compared to all other models.

From the Table 3.4 it is found that the covariates AGE, SHO, CHF are showing significant differences of the event of interest in all models. The deviance is very high in Exponential AFT model than in Weibull AFT, Lognormal AFT and Log-logistic AFT. Among the models, Weibull AFT model has the lowest level of deviance compared to all other models. The PH model displays significant lack of fit while the AFT model describes the data well. The AFT model results are smaller deviance than the PH model. From a clinical trial the AFT model in applications have seem to be more appropriate modeling and has the added advantage of being easier to interpret. It is found that the AFT model
should be considered as an alternative to the PH model in the analysis of time to event data.

In clinical trial applications the AFT model is often a more realistic model than the PH model in the analysis of time to event data. The PH model is appropriate when there is a difference between the groups in the longer term in the context of the follow-up period. The AFT model is more appropriate when the group differences are seen over a shorter timeframe while in the longer term the probability of remaining event free is similar in the two groups. It is argued that PH model is not always appropriate and that the AFT model in many applications provides a more appropriate modeling framework and has the added advantage of being straightforward to interpret than the proportional hazards model.