

## ON FRAILTY MODELS FOR KIDNEY INFECTION DATA WITH EXPONENTIAL BASELINE DISTRIBUTION

S. G. PAREKH<sup>1</sup>, D. K. GHOSH<sup>2</sup> & S. R. PATEL<sup>3</sup>

<sup>1</sup>Faculty of Business Administration, Dharmsinh Desai University, Nadiad, India

<sup>2</sup>Department of Statistics, Saurashtra University, Rajkot, India

<sup>3</sup>Department of Statistics, Sardar Patel University, Vallabh Vidyanagar, India

### ABSTRACT

The paper deals with the estimation of survival function with the use of linear hazard function and exponential base line distribution. Considering cox PH model with exponential base line distribution, the maximum likelihood estimator of general linear parametric function of regressors and frailty parameter has been obtained by taking very general form of the regression matrix and estimator being biased its MSE is obtained. Further the same has been estimated by least square theory using Gauss-Markoff model and it has been compared with usual estimator. The results have been applied by taking kidney infection data.

**KEYWORDS:** AN, BLUE, Frailty Models, GN, MLE, PKD

### 1. INTRODUCTION

The Cox proportional hazard model is extensively used in medical research as Survivor model. Cox proportional hazard model uses regression analysis for censored data. In Cox proportional hazard model the explanatory variables or covariates are studied for the random effect of covariates on distribution of survivor times. Unfortunately we are unable to include the relevant covariates related to diseases in many cases due to unawareness of the factor.

When random effects (frailties) are included, the failure time distribution obtained by integration over the distribution of frailties loses this simple cancellation property. For such failure time distribution Hougaard (1986a, 1986b), Prentice, Williams and Peterson (1981) have introduced some models. The Maximum Likelihood Estimation (MLE) method and the Best Linear Unbiased Predictor (BLUP) method have been used by Handerson (1975) and McGilchrist and Aisbett(1991).

Particularly useful family of baseline model is obtained from a univariate lifetime model with hazard function  $h_0(t)$  by defining

$$h(t/x) = h_0(t)e^{\underline{X}'\underline{\beta}}$$

where  $\underline{X} = (x_1, x_2, \dots, x_p)'$  and  $\underline{\beta} = (\beta_1, \beta_2, \dots, \beta_p)'$

and by including the baseline distribution of  $\underline{z}$ , the hazard function will be

$$h(t/x) = z \cdot h_0(t)e^{\underline{X}'\underline{\beta}}$$

In this article we consider the baseline distribution of  $z$  as exponential distribution defined as

$$f(t) = \theta^{-1} e^{-\frac{t}{\theta}}, \quad t \geq 0, \theta > 0$$

Let  $T$  be the random variable associated with the survivor function,  $S(t)$ . One would like to examine the effect of particular regressor variables in relation to the proportional hazard assumption such as  $Y = \log T = \underline{X}'\underline{\beta} + \Psi\theta$  taking  $\Psi\theta = \mathbf{e}_i$ ,  $\mathbf{e}_i$  being error term which corresponds to shared frailty model which relates to the covariates in addition to regressor covariates  $\underline{X}$ . Since the exponential distribution has hazard function constant and whenever hazard function turns out to be constant, the base line distribution is exponential distribution and therefore we consider the base line distribution of each  $\mathbf{e}_i$  as

$$f(\mathbf{e}_i) = \theta^{-1} e^{-\frac{\mathbf{e}_i}{\theta}}, \quad \mathbf{e}_i \geq 0, \theta > 0, \quad i = 1, 2, \dots, n$$

with  $E(\mathbf{e}_i) = \theta$  and  $Var(\mathbf{e}_i) = \theta^2$ ,  $Cov(\mathbf{e}_i, \mathbf{e}_j) = 0$ , for  $i \neq j$ ,  $i = 1, 2, \dots, n$ ,  $j = 1, 2, 3, \dots, n$

So the model for  $X$  turns out as

$$\underline{Y} = \underline{X}\underline{\beta} + \underline{\mathbf{e}} \quad (1.1)$$

where  $\underline{Y}$  is  $n \times 1$  vector,  $(Y_1, Y_2, \dots, Y_n)$ ,  $X$  is  $n \times p$  matrix  $(X_{ij}, i = 1, 2, \dots, n; j = 1, 2, \dots, p)$  of known constants,  $\underline{\beta}$  is  $p \times 1$  vector of parameters  $\beta_1, \beta_2, \dots, \beta_p$  and  $\underline{\mathbf{e}} = (\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_p)'$  is error vector responsible for frailty model. We assume that the distribution of each component  $\mathbf{e}_i$  of  $\underline{\mathbf{e}}$  is to be referred as baseline exponential distribution.

In this article we use Maximum Likelihood (ML) and Least square methods for estimating linear function of regression coefficients and frailty parameter such as

$$\Delta = \underline{C}'\underline{\beta} + \delta\theta$$

where  $\underline{C} = (C_1, C_2, \dots, C_p)'$  and  $\delta$  are known, which relates to the frailty and shared frailty parameters  $\underline{\beta}$  and  $\theta$  in section 2. By using Gauss-Marcoff model in Cox PH model the estimation of regression and frailty parameter by taking exponential base line distribution has been dealt with in section 3. Further both methods have been illustrated by taking real data obtained by McGilchrist and Aisbett (1991) on kidney infection in section 4.

## 2. ESTIMATION OF $\Delta$ BY MAXIMUM LIKELIHOOD METHOD.

In Cox PH model, let individual frailty variable  $\mathbf{e}_i$  follow exponential distribution with p.d.f. as

$$f(\mathbf{e}_i) = \theta^{-1} e^{-\frac{\mathbf{e}_i}{\theta}}, \quad \mathbf{e}_i \geq 0, \theta > 0, \quad i = 1, 2, \dots, n \quad (2.1)$$

with assumptions given in section 1.

Let  $Y_i = \log T_i$  where  $T_i$  is the time recorded for the  $i^{th}$  individual and  $Y_i$  follows negative exponential distribution with Cox PH model, having  $e^{-\underline{X}_i' \underline{\beta}}$ .

where  $\underline{X}_i' = (X_{i1}, X_{i2}, \dots, X_{ip})$  as

$$f_{Y_i}(y) = \frac{1}{\theta} \cdot \exp \left[ -\frac{(y_i - \underline{X}_i' \underline{\beta})}{\theta} \right], \quad y \geq \underline{X}_i' \underline{\beta}, \theta > 0$$

Now considering  $\underline{\epsilon} = (\epsilon_1, \epsilon_2, \dots, \epsilon_p)$  as parameters of n-variate negative exponential distribution the model given in (1.1) is

$$\underline{Y} = \underline{X}\underline{\beta} + \underline{\epsilon}.$$

Using the mean and variance of the exponential distribution of  $\epsilon_i$ , we have

$$E(\underline{\epsilon}) = \theta \underline{1} \quad \text{where } \underline{1} \text{ is n-vector of unity.}$$

$$E(\underline{Y}) = \underline{X}\underline{\beta} + \theta \underline{1} \text{ and } D(\underline{\epsilon}) = D(\underline{Y}) = \theta^2 I$$

Let  $Y_1, Y_2, \dots, Y_n$  be a random sample on  $Y$  and let  $Y_{(1)}, Y_{(2)}, \dots, Y_{(n)}$  be the corresponding ordered random sample. Thus the likelihood  $L(\underline{\beta}, \theta)$  of  $Y_1, Y_2, \dots, Y_n$  at  $y_1, y_2, \dots, y_n$  being

$$\begin{aligned} L(\underline{\beta}, \theta) &= \theta^{-n} \exp \left( -\sum_{i=1}^n (y_i - \underline{x}'_i \underline{\beta}) / \theta \right) \\ &= \theta^{-n} \exp \left( -\underline{1}'(\underline{y} - \underline{X}\underline{\beta}) / \theta \right) \end{aligned}$$

the m.l.e. of  $\underline{\beta}$  of  $\underline{\beta}$  is given by

$$\underline{X}\underline{\hat{\beta}} = \underline{1}Y_{(1)}.$$

$$\text{i.e. } \underline{X}'\underline{X}\underline{\hat{\beta}} = \underline{X}'\underline{1}Y_{(1)}$$

$$\Rightarrow \underline{\hat{\beta}} = \begin{cases} (\underline{X}'\underline{X})^{-1} \underline{X}'\underline{1}Y_{(1)}, & \text{if } |\underline{X}'\underline{X}| \neq 0 \\ (\underline{X}'\underline{X})^{-1} \underline{X}'\underline{1}Y_{(1)}, & \text{if } |\underline{X}'\underline{X}| = 0. \end{cases} \tag{2.2}$$

Further,

$$\frac{\partial \log L(\underline{\hat{\beta}}, \theta)}{\partial \theta} = 0$$

leads to the m.l.e.  $\hat{\theta}$  of  $\theta$  as

$$\hat{\theta} = \bar{Y} - Y_{(1)}, \quad (2.3)$$

where

$$\bar{Y} = \sum_{i=1}^n \frac{Y_i}{n}$$

For given p-vector  $\underline{C}$  and a given constant  $\delta$ , the m.l.e.  $\hat{\Delta}$  of  $\Delta = \underline{C}'\underline{\beta} + \delta\theta$  is

$$\hat{\Delta} = \mathbf{a}Y_{(1)} + \delta\bar{Y}, \quad (2.4)$$

where

$$\mathbf{a} = \begin{cases} \underline{C}'(X'X)^{-1}X'\underline{1} - \delta, & \text{if } |X'X| \neq 0 \\ \underline{C}'(X'X)^-X'\underline{1} - \delta, & \text{if } |X'X| = 0 \end{cases} \quad (2.5)$$

where  $(X'X)^-$  is generalised inverse of matrix  $X'X$  in the sense of  $AA^-A = A$ .

If  $\underline{C}$  belongs to the row space of  $X'X$  and  $\underline{1}$  belongs to the column space of  $X'X$  then  $\hat{\Delta}$  is obviously unique and unbiased.

By using

$$E(\bar{Y}) = \theta + \underline{X}_i'\underline{\beta} \text{ and } E(Y_{(1)}) = \frac{\theta}{n} + \underline{X}_i'\underline{\beta}.$$

$$\hat{\Delta} = \underline{C}'\underline{\beta} + \delta\hat{\theta}$$

$$= (\underline{C}'(X'X)^{-1}X'\underline{1} - \delta)Y_{(1)} + \delta\bar{Y}$$

$$E(\hat{\Delta}) = [\underline{C}'(X'X)^{-1}X'\underline{1} + (n-1)\delta]\frac{\theta}{n} + \underline{C}'\underline{\beta}$$

$$= (\mathbf{a} + n\delta)\frac{\theta}{n} + \underline{C}'\underline{\beta} \quad \text{where } \mathbf{a} \text{ is defined in (2.5)}$$

$\hat{\Delta}$  is unbiased for  $\Delta$  if  $\underline{C}$  belongs to the row space of  $X'X$  and  $\underline{1}$  belongs to the column space of  $X'X$ .

Further, also  $\hat{\Delta}$  is unbiased for  $\Delta$  if

$$\begin{cases} \underline{C}'(X'X)^{-1}X'\underline{1} = \delta, & \text{if } |X'X| \neq 0 \\ \underline{C}'(X'X)^-X'\underline{1} = \delta, & \text{if } |X'X| = 0 \end{cases} \quad (2.6)$$

otherwise it is biased and the bias is

$$\text{Bias} = [\underline{C}'(X'X)^{-1}X'\underline{1} - \delta]\frac{\theta}{n}$$

$$= \frac{\mathbf{a}\theta}{n}$$

If  $\hat{\Delta}$  is unbiased under condition (2.6), the variance of the m.l.e.  $\hat{\Delta}$  of  $\Delta$  is

$$\begin{aligned} \text{Var}(\hat{\Delta}) &= E\{\hat{\Delta} - E(\hat{\Delta})\}^2 \\ &= \mathbf{a}^2 \text{Var}(Y_{(1)}) + \delta^2 \text{Var}(\bar{Y}) + 2\mathbf{a}\delta \text{Cov}(Y_{(1)}, \bar{Y}) \\ &= \frac{\mathbf{a}^2 \theta^2}{n} + \frac{\delta^2 \theta^2}{n} + \frac{2\mathbf{a}\delta \theta^2}{n} \quad \text{where } \mathbf{a} \text{ is defined in (2.2.5)} \\ &= (\mathbf{C}'(X'X)^{-1}X'\mathbf{1})^2 \frac{\theta^2}{n} \end{aligned} \tag{2.7}$$

and the estimated  $\text{Var}(\hat{\Delta})$  is

$$\widehat{\text{Var}}(\hat{\Delta}) = (\mathbf{C}'(X'X)^{-1}X'\mathbf{1})^2 \frac{\hat{\theta}^2}{n}.$$

Also the estimated standard error is

$$\text{s. e.}(\hat{\Delta}) = \sqrt{\widehat{\text{Var}}(\hat{\Delta})} \tag{2.8}$$

The variance of the unbiased estimator of  $\Delta$  under the conditions of estimability either given in (2.6) or  $\underline{\mathbf{C}}$  belongs to the row space of  $X'X$  and  $\mathbf{1}$  belongs to the column space of  $X'X$ , is  $\frac{\delta^2 \theta^2}{n}$ , which does not involve  $\underline{\beta}$ . We note that for fixed  $\underline{\mathbf{C}}, \delta, n, p$  the minimization of  $\text{Var}(\hat{\Delta})$  is equivalent to the minimization of  $\frac{\delta^2 \theta^2}{n}$  when  $\mathbf{C}'(X'X)^{-1}X'\mathbf{1} = \delta$  holds.

When the m.l.e.,  $\hat{\Delta}$  of  $\Delta$  is not unbiased estimator of  $\Delta$ , the Mean Square Error(MSE) of  $\hat{\Delta}$  is

$$\begin{aligned} \text{MSE}(\hat{\Delta}) &= E(\hat{\Delta} - \Delta)^2 \\ &= E(\hat{\Delta} - E\hat{\Delta} + E\hat{\Delta} - \Delta)^2 \\ &= E(\hat{\Delta} - E\hat{\Delta})^2 + (E\hat{\Delta} - \Delta)^2 + 2E(\hat{\Delta} - E\hat{\Delta})(E\hat{\Delta} - \Delta) \\ &= \frac{\mathbf{a}^2 \theta^2}{n^2} + \frac{\delta^2 \theta^2}{n} + \frac{2\mathbf{a}\delta \theta^2}{n} + \left(\frac{\mathbf{a}\theta}{n}\right)^2 \\ &= \text{Var}(\hat{\Delta}) + \left(\frac{\mathbf{a}\theta}{n}\right)^2 \end{aligned} \tag{2.9}$$

If  $\frac{\mathbf{a}\theta}{n} \geq 0$ , the unbiased estimator of  $\Delta$  is better than MSE estimator of  $\Delta$ .

As  $\theta > 0, n > 0$ , the condition  $\frac{\mathbf{a}\theta}{n} (\geq 0)$  reduces to  $\mathbf{a} (\geq 0)$

$$\text{i.e. } C'(X'X)^{-1}X'\underline{\mathbf{1}} \geq \delta \quad (2.10)$$

Generally,  $C'(X'X)^{-1}X'\underline{\mathbf{1}}$  is less than 1 and  $\delta$  is generally taken as 1. In general MSE estimator is better than unbiased estimator when Cox PH frailty model is used.

### 3. ESTIMATION OF $\Delta$ BY LEAST SQUARE METHOD OF SHARED EXPONENTIAL FRAILTY MODEL

Since the p.d.f  $f(t_i)$  of survivor time  $T_i$  of  $i^{\text{th}}$  individual is expressed in shared frailty model as

$$f(t_i) = h_0(t_i) e^{(x_i\beta_i + w_i\Psi_i)} \quad (3.1)$$

where  $w_i\Psi_i$  is shared frailty variable and  $h_0(t_i)$  is hazard function of exponential distribution which is constant,  $\alpha$  and the p.d.f. of  $n$  variates  $T_1, T_2, \dots, T_n$  will be

$$L = \prod_{i=1}^n f(t_i) = e^{(\sum x_i\beta_i + \sum w_i\Psi_i)}$$

Taking logarithm, we get

$$\begin{aligned} Y &= \sum \log f(t_i) \\ &= \sum x_i\beta_i + \sum w_i\Psi_i \\ &= \underline{X}'\underline{\beta} + \underline{W}'\underline{\Psi} \end{aligned} \quad (3.2)$$

and denoting  $\underline{W}'\underline{\Psi}$  by  $\underline{e}$  and taking  $n$ -vector  $\underline{Y}$  as  $n$ -variate frailty exponential distribution the model is

$$\underline{Y} = \underline{X}\underline{\beta} + \underline{e}$$

which is defined in (1.1).

Now reconstructing the model (1.1) as

$$\underline{Y} = \underline{X}\underline{\beta} + \underline{\theta}\underline{\mathbf{1}} + \underline{e}^* \quad (3.3)$$

where  $\underline{e}^* = \underline{e} - \underline{\theta}\underline{\mathbf{1}}$  such that  $E(\underline{e}^*) = 0$  and  $D(\underline{e}^*) = \theta^2 I$

$$\text{Let } \phi(\underline{\beta}, \theta) = (\underline{Y} - \underline{X}\underline{\beta} - \underline{\theta}\underline{\mathbf{1}})'(\underline{Y} - \underline{X}\underline{\beta} - \underline{\theta}\underline{\mathbf{1}})$$

then least square method demands to solve the following equations

$$\frac{\partial \phi(\underline{\beta}, \theta)}{\partial \underline{\beta}} = \underline{0}$$

and

$$\frac{\partial \phi(\underline{\beta}, \theta)}{\partial \theta} = 0$$

Now

$$\frac{\partial \phi(\underline{\beta}, \theta)}{\partial \underline{\beta}} = \underline{0} \Rightarrow X'X\underline{\beta} = X'\underline{Y} - X'\underline{\theta}\underline{1} \tag{3.4}$$

$$\frac{\partial \phi(\underline{\beta}, \theta)}{\partial \theta} = \underline{0} \Rightarrow \underline{1}'X\underline{\beta} + \underline{1}'\underline{1}\theta = \underline{1}'\underline{Y} \tag{3.5}$$

From (3.4) we have  $\hat{\underline{\beta}}$  in terms of  $\theta$  as

$$\begin{aligned} \hat{\underline{\beta}} &= \begin{cases} (X'X)^{-1}X'\underline{Y} - (X'X)^{-1}X'\underline{1}\theta & \text{if } |X'X| \neq 0 \\ (X'X)^-X'\underline{Y} - (X'X)^-X'\underline{1}\theta & \text{if } |X'X| = 0 \end{cases} \\ &= M(\underline{Y} - \underline{1}\theta) \end{aligned} \tag{3.6}$$

where  $M = (X'X)^{-1}X'$ , taking  $|X'X| \neq 0$

Substituting  $\hat{\underline{\beta}}$  in (3.5), we get

$$\begin{aligned} \underline{1}'X[M(\underline{Y} - \underline{1}\theta)] + \underline{1}'\underline{1}\theta &= \underline{1}'\underline{Y} \\ \Rightarrow \underline{1}'XMY - \underline{1}'XM\underline{1}\theta + \underline{1}'\underline{1}\theta &= \underline{1}'\underline{Y} \\ \Rightarrow [\underline{1}'\underline{1} - \underline{1}'XM\underline{1}]\theta &= \underline{1}'\underline{Y} - \underline{1}'XMY \\ \Rightarrow \underline{1}'[I - XM]\underline{1}\theta &= \underline{1}'[I - XM]\underline{Y} \\ \Rightarrow \underline{1}'[I - X(X'X)^{-1}X']\underline{1}\theta &= \underline{1}'[I - X(X'X)^{-1}X']\underline{Y} \\ \Rightarrow \underline{1}'A\underline{1}\theta &= \underline{1}'A\underline{Y}, \quad \text{where } A = I - X(X'X)^{-1}X' \text{ and is} \end{aligned}$$

idempotent matrix.

$$\Rightarrow \hat{\theta} = \frac{\underline{1}'A\underline{Y}}{\underline{1}'A\underline{1}} \tag{3.7}$$

Using (3.7) in (3.6)

$$\hat{\underline{\beta}} = \begin{cases} (X'X)^{-1}X'\underline{Y} - (X'X)^{-1}X'\underline{1} \frac{\underline{1}'A\underline{Y}}{\underline{1}'A\underline{1}} & \text{if } |X'X| \neq 0 \\ (X'X)^-X'\underline{Y} - (X'X)^-X'\underline{1} \frac{\underline{1}'A\underline{Y}}{\underline{1}'A\underline{1}} & \text{if } |X'X| = 0 \end{cases} \tag{3.8}$$

Also,

$$D(\hat{\underline{\beta}}) = D(X'X)^{-1}X'(\underline{Y} - \underline{\theta}\underline{1})$$

$$\begin{aligned}
 &= (X'X)^{-1}X'D(\underline{Y} - \underline{\theta 1})X(X'X)^{-1} \\
 &= (X'X)^{-1}\theta^2
 \end{aligned}
 \tag{3.9}$$

**REMARK (3.1)**

We note that  $(X'X)^{-1}X'Y$  is the usual regression estimator of  $\underline{\beta}$ . Due to shared frailty, the estimator is reduced by the quantity  $(X'X)^{-1}X'1 \frac{1'AY}{1'A1}$  which shows that the frailty has negative effect in the estimation of primary covariates.

Particularly, when  $\underline{\theta} = \bar{Y}$ , the shared frailty is neglected and in this case the regression frailty is usual, regression estimator,  $\underline{\beta} = (X'X)^{-1}X'Y$ , and hence no new covariates are necessary to explain hazard rate or survivor function. This is due to exponential base line distribution.

**4. APPLICATION TO KIDNEY INFECTION PATIENTS.**

For illustration we use the kidney infection data given by McGilchrist and Aisbett (1991) of 38 kidney Patients using portable dialysis. A catheter is inserted to each admitted kidney Patient and keeps it until infection is removed, that is infection is cleared. After some time again catheter is inserted and removed when infection occurs due to kidney or some other reasons. Time (in number of days) recorded when catheter is reinserted. The purpose is to note the incidence rate of infection to risk variables (Age, Sex and disease type of kidney disorder, Glomerulo Neptiritis(GN), Acute Neptiritis(AN) and Polycyatic Neptiritis(PKD). The time interval recorded is considered as random variable and according to Cox PH model we have illustrated the estimation by MLE, using the best linear unbiased prediction (BLUP). Estimation of the problem is described in McGilchrist and Aisbett (1991). We have obtained maximum likelihood estimator (M.L.E.) of main covariates Age( $X_1$ ), Sex( $X_2$ ) along with other correlated covariates like GN( $X_3$ ), AN( $X_4$ ) and PKD( $X_5$ ). Also BLUP of regression coefficient,  $\beta_1, \beta_2, \beta_3, \beta_4, \beta_5$  and  $\theta$  with its linear function  $\Delta = \underline{c}' \underline{\beta} + \delta \theta$  is obtained. Not only by MLE Method linked with BLUP but obtained independently using linear model.

We give below the estimates of regression coefficients,  $\beta_1, \beta_2, \beta_3, \beta_4, \beta_5$  and  $\theta$  on the basis of the data given in McGilchrist and Aisbett (1991) by taking time variate  $T$  of recording time when catheter is inserted first time and  $Y$  is  $\log T$ . The maximum likelihood estimates are obtained by using (2.2), (2.3) and (2.8). where  $X_{38 \times 5}$  is a matrix representing 38 patients and 5 covariates  $X_1, X_2, X_3, X_4, X_5$  and  $Y_{(1)} = \min(Y_i, i = 1, 2, \dots, 38), Y_i = \log T_i$

**4.1. APPLICATION OF (2.2), (2.3) AND (2.8) TO GET M.L.E. OF  $\beta_1, \beta_2, \beta_3, \beta_4, \beta_5$  AND  $\theta$**

Following Table 4.1 gives the estimates of  $\beta_1, \beta_2, \beta_3, \beta_4, \beta_5$  by M.L. method and their standard errors.

**Table 4.1**

Variable	Age	Sex	GN	AN	PKD
Regression coefficient estimate	0.011535	0.228179	0.002059	-0.08724	0.044094
Standard error	0.008532	0.168784	0.001523	0.064528	0.032616

The estimate of the variance is 0.021587. In general the effect of the prior distribution on frailty is to shrink the estimate towards the origin. The only regression coefficients i.e. significantly large compared to its standard error is that of the age variable, indicating a lower infection rate for the lower age of the patients.

$$\hat{\theta} = \bar{Y} - Y_{(1)} = 3.160408$$

The value of frailty parameter is

$$\ln(\hat{\theta}) = \ln(3.160408) = 1.150701$$

We note that taking  $C_1 = C_2 = C_3 = C_4 = C_5 = 1$  and  $\delta = 1$  that is,  $C' \underline{\beta} + \delta \theta = \beta_1 + \beta_2 + \dots + \beta_5 + \theta$ , the MSE of  $C' \underline{\hat{\beta}} + \delta \hat{\theta}$  turns out as 1.349332 and the condition given in (2.10) is not justified and hence MSE estimator is better than unbiased estimator of frailty parameters.

**Particular Cases**

S.N.	Cases						$\hat{\Delta}$	s.e( $\hat{\Delta}$ )
	$C_1$	$C_2$	$C_3$	$C_4$	$C_5$			
1	1	0	0	0	0	1.162236	0.008532	
2	0	1	0	0	0	1.37888	0.168784	
3	1	1	1	1	1	1.349332	0.146927	
4	1	1	0	0	0	1.390415	0.177316	

This concludes the combined effect of age and sex on the estimate being significantly large. Age and sex has lower infection rate. Thus the primary covariates are much more significant rather than other covariates.

**4.2. APPLICATION OF (3.8) AND (3.9) TO GET BLUP OF  $\beta_1, \beta_2, \beta_3, \beta_4, \beta_5$  AND  $\theta$ .**

Using the result (3.8) and (3.9) the usual estimates of  $\beta_1, \beta_2, \beta_3, \beta_4, \beta_5$  and their s.e. are presented in table 4.2. Also the BLUP estimates of  $\beta_1, \beta_2, \beta_3, \beta_4, \beta_5$  with their s.e. are presented in table 4.3.

Here, we have , 
$$\hat{\theta} = \frac{1'AY}{1'A1} = 3.60439$$

The value of frailty parameter is  $\ln(\hat{\theta}) = \ln(3.60439) = 1.282152$

**Table 4.2: Usual Estimates (Without Frailty) of  $\beta_1, \beta_2, \beta_3, \beta_4, \beta_5$  with s.e**

Variable	Age( $\beta_1$ )	Sex( $\beta_2$ )	GN( $\beta_3$ )	AN( $\beta_4$ )	PKD( $\beta_5$ )
Regression coefficient estimate	0.049381	2.68169	-0.4497	-1.4122	0.364796
Standard error	0.014662	0.520854	0.768717	0.759195	0.979725

**Table 4.3: Regression Estimates (With Frailty) of  $\beta_1, \beta_2, \beta_3, \beta_4, \beta_5$  with s.e**

Variable	Age( $\beta_1$ )	Sex( $\beta_2$ )	GN( $\beta_3$ )	AN( $\beta_4$ )	PKD( $\beta_5$ )
Regression coefficient estimate	0.028043	2.259586	-0.45351	-1.25082	0.283228
Standard error	0.014662	0.520854	0.768717	0.759195	0.979725

From above two tables (4.2) and (4.3) we observe that the estimates of frailty parameters are smaller than those of usual BLUP values of  $\beta_1, \beta_2, \beta_3, \beta_4, \beta_5$  even though their estimates of s.e. are same.

## CONCLUSIONS

Theoretically it is observed that the m.l.e. of primary covariates when frailty covariates are present in terms of hazard function parameter with cox PH model are more advisable to use rather than usual BLUP estimator and this is supported by data on kidney infection.

## REFERENCES

1. Cox, D. R. (1972). Regression models and life tables (with discussion). *Journal of the Royal Statistical Society, Series B* 34, 187-200.
2. Handerson, C. (1975). Best linear unbiased estimation and prediction under a selection model. *Biometrika*, 31, 423-447.
3. Hougaard, P. (1986a). Survival models for heterogeneous populations derived from stable distribution. *Biometrika*, 73, 387-396.
4. Hougaard, P. (1986b). A class of multivariate failure time distributions. *Biometrika*, 73, 671-678.
5. McGilchrist, C. A. and Aisbett, C.W. (1991). Regression with frailty in survival analysis. *Biometrics*, 47, 461-466.
6. Prentice, R. P. (1981). On regression analysis of multivariate failure time data. *Biometrika*, 68, 373-379.