International Journal of Computational and Theoretical Statistics
ISSN (2210-1519)
Int. J. Comp. Theo. Stat. 5, No. 1 (May-2018)

http://dx.doi.org/10.12785/ijcts/050104

Some Univariate Continuous Frailty Models

S. G. Parekh [1] and S. R. Patel[2]

[1] *Faculty of Business Administration, Dharmsinh Desai University, Nadiad, India*
[2] *Department of Statistics, Sardar Patel University, Vallabh Vidyanagar, India*

Submitted Nov 5, 2017, Revised January 16, 2018, Accepted Feb 9, 2018, Published May 1, 2018

Abstract: In this paper we have considered some univariate distributions and their frailty models, such as Gamma frailty model, Weibull frailty model, generalized exponential frailty model, Exponential power distribution as frailty model, Log normal frailty model. In all these frailty distributions we have obtained maximum likelihood estimates of the parameters of the simulated data and used M. C. Method.

Keywords: Frailty Models, Gamma, Weibull, Generalized exponential, Exponential power, Log normal frailty model, m.l.e., simulation

1. INTRODUCTION

Parekh et.al. (2016) have considered discrete frailty models, we extended the frailty models for different continuous distributions such as Gamma, Weibull, Generalized exponential, Exponential power and Log normal frailty models. Section 2 deals with Gamma frailty model in which we have derived m.l.e. of the parameters of simulated data. In section 3 we considered Weibull frailty model. Section 4 discusses Generalized exponential frailty model. Section 5 deals with Exponential power distribution as frailty model and in section 6 we considered Log normal frailty model.

2. GAMMA FRAILTY MODEL

The gamma distribution has been widely applied as a mixture distribution (e.g. Vaupel et al., 1979; Congdon, 1995; Hougaard, 2000). From a computational and analytical point of view, it fits very well to failure data because it is easy to derive the closed form expressions of unconditional survival, cumulative density and hazard function. This is due to the simplicity of the Laplace transform and due to its uses in most applications. It is a flexible distribution that takes a variety of shapes as $k$varies such as when k= 1, it turns out to exponential distribution and when kis large, it takes a normal distribution. The use of gamma distributions for frailties in time-to-event data analysis is suggested by Abbring and van den Berg (2005). The gamma distribution under mild regularity assumptions, a large class of frailty model converges to it.

As frailty cannot be negative and the gamma distribution and also the log-normal distribution is the most commonly used distributions to model variables that are necessarily positive. Furthermore, it turns out that the assumption that frailty at the beginning of the follow-up is gamma distributed yields some useful mathematical results. This includes the following

(1) Frailty among the survivors at any time t is gamma distributed with the same value of the shape parameter k as at birth or at beginning of follow-up. The value of the second parameter, however, is now given by $\lambda + H_0(t)$, where $H_0(t)$ denotes the cumulative baseline hazard function.

(2) Frailty among those who die at any age t is also gamma distributed, with the same parameter $\lambda + H_0(t)$as among those surviving to t but with shape parameter (k+1). In particular, it follows that the mean frailty among the deaths at age t is $\frac{\lambda+1}{\lambda+H(t)}$ compared to $\frac{\lambda}{\lambda+H(t)}$among the survivors at the same age. This demonstrates the selection by death of the high risk individuals, e.g. the individuals with high values of frailty Z. When k = λ in the gamma distribution EZ = 1 and $\frac{1}{\lambda} = \sigma^2$, the variance of the frailty variable.

Considering X has frailty distribution as $G(\lambda, k)$ where the p.d.f. of $G(\lambda, k)$ is

E-mail address: sg_parekh@yahool.com, dipa_parekh@yahoo.co.in

$G(t;\lambda,k)=\frac{\lambda^k}{\Gamma k}t^{k-1}e^{-\lambda t},\ t>0,\ \lambda>0,\ k>0$

and Laplace transform is

$$\begin{aligned} L(s) &= Ee^{-ST} \\ &= \frac{\lambda^k}{\Gamma k}\int_0^\infty e^{-st}\,t^{k-1}e^{-\lambda t}dt \\ &= \frac{\lambda^k}{\Gamma k}\int_0^\infty e^{-(s+\lambda)t}\,t^{k-1}dt \\ &= \frac{\lambda^k}{\Gamma k}\cdot\frac{\Gamma k}{(s+\lambda)^k} \\ &= \left(\frac{\lambda}{s+\lambda}\right)^k \\ &= \left(\frac{s+\lambda}{\lambda}\right)^{-k} \\ &= \left(1+\frac{s}{\lambda}\right)^{-k} \end{aligned}$$

Now,

$$\begin{aligned} S(t) &= e^{-H(t)} \\ &= L(H(t)) \\ \therefore L(H_0(t)) &= \left(1+\frac{H_0(t)}{\lambda}\right)^{-k} \end{aligned}$$

Since, $k=\lambda=\frac{1}{\sigma^2}$

Therefore the unconditional survival function and hazard function are

$$S(t) = L\big(H_0(t)\big) = (1+\sigma^2 H_0(t))^{-\frac{1}{\sigma^2}} \text{ and } h(t)=\frac{h_0(t)}{1+\sigma^2 H_0(t)} \tag{2.1}$$

There are many applications of the gamma frailty model. In a very early paper Lancaster (1979) suggested this model for the duration of unemployment spells and Vaupel et al. (1979) used it to correct life tables of heterogeneous populations. Aalen (1987) studied the expulsion of intra-uterine contraceptive devices. Manton et al. (1981) used it for comparing the mortality experience of heterogeneous populations, Manton and Stallard (1981) to explain the black/white mortality crossover in the US, Manton et al. (1986) compared the inverse normal and the gamma models, together with Gompertz and Weibull baseline hazards, in a study of survival at advanced ages, based on Medicare data. Jeong et al. (2003) used a gamma frailty to model long-term follow-up survival data from breast cancer clinical trials when the treatment effect diminishes over time as an alternative to the proportional hazards model. Jones (1998) used a gamma-Gompertz model for analyzing the impact of selective lapsation on mortality in life insurance.

2.1 SIMULATION STUDY OF FRAILTY GAMMA DISTRIBUTION

To evaluate the frailty model of gamma distribution we carried out a simulation study of 5000 observations simulated from gamma distribution with p.d.f.

$$f(t)=\frac{1}{\Gamma\alpha}t^{\alpha-1}e^{-t},\ \ t>0,\alpha>0 \tag{2.2}$$

Taking initial value $\alpha=6$

2.2 MAXIMUM LIKELIHOOD ESTIMATOR OF PARAMETER α OF GAMMA DISTRIBUTION

Likelihood function of simulated values $t_1,t_2,\dots,t_{5000}$ of gamma distribution is

$$\begin{aligned} L &= \left(\frac{1}{\Gamma\alpha}\right)^n \prod_{i=1}^n ({t_i}^{\alpha-1})\,e^{-\sum t_i} \\ &= \frac{\tilde{t}^{n(\alpha-1)}e^{-n\bar{t}}}{(\Gamma\alpha)^n}, \text{ where } \bar{t}=\frac{\sum t_i}{n},\ \tilde{t}=(\textstyle\prod_{i=1}^n t_i)^{\frac{1}{n}} \\ \therefore logL &= -nlog\Gamma\alpha+n(\alpha-1)log\tilde{t}-n\bar{t} \\ \frac{\partial logL}{\partial\alpha} &= \frac{-n\frac{\partial\Gamma\alpha}{\partial\alpha}}{\Gamma\alpha}+nlog\tilde{t}=0 \text{ gives} \end{aligned}$$

$$log\tilde{t}-\frac{\partial log\Gamma\alpha}{\partial\alpha}=0 \tag{2.3}$$

We have to solve (2.3) for α

Taking $\frac{\partial log\Gamma\alpha}{\partial\alpha}=\psi(\alpha)$, (2.3) will be

$$g(\alpha)=log\tilde{t}-\psi(\alpha) \tag{2.4}$$

We have evaluated (2.4) at different values of α as under

The following table 2.1 shows the estimates of parameter α and its standard error.

TABLE 1. MLE FOR SIMULATION STUDY OF GAMMA DISTRIBUTION

Initial value of α	$\hat{\alpha}$	**Var($\hat{\alpha}$)**	**S.E.($\hat{\alpha}$)**	**converge**	**llvalue**	**no. of iterations**
5.90000	5.855540	0.001074192	0.03277487	0	11230.00	14
6.00000	5.972765	0.001097572	0.03312962	0	11322.26	16
6.02510	6.090334	0.001121023	0.03348169	0	11346.89	15
6.03515	6.036795	0.001110344	0.03332182	0	11288.42	14
6.03525	5.971532	0.001097326	0.03312591	0	11253.46	14
6.04000	6.100373	0.001123026	0.03351158	0	11280.21	16
6.05000	6.032300	0.001109447	0.03330836	0	11357.72	16

llvalue = value of the loglikelihood.

From the table (2.1) it is observed that value of alpha.est. decreases at α = 6.02510, therefore m.l.e. $\hat{\alpha}$ = 6.02510 with s.e.($\hat{\alpha}$) = 0.03348169

3. WEIBULL FRAILTY MODEL

The p.d.f.ofweibull distribution is

$$f(t;\alpha,\lambda) = \alpha\lambda t^{\alpha-1}e^{-\lambda t^{\alpha}}, t > 0, \alpha > 0, \lambda > 0$$

We generated 5000 observations $t_1, t_2, \dots, t_{5000}$ using "R" from above distribution with initial values of $\alpha = 2.0$ and $\lambda = 4.2$. The log likelihood of the observations is

$$logL = nlog\alpha + nlog\lambda + (\alpha - 1)\sum logt_i - \lambda \sum t_i{}^{\alpha}$$

$$\frac{\partial logL}{\partial \alpha} = \frac{n}{\alpha} + \sum logt_i - \lambda \sum t_i{}^{\alpha} logt_i = 0$$

Taking $logt_i = y_i$, $t_i = e^{y_i}$, we have likelihood equation as

$$\frac{1}{\alpha} + \frac{\sum y_i}{n} - \frac{\lambda}{n}\sum y_i e^{\alpha y_i} = 0 \tag{3.1}$$

and $\frac{\partial logL}{\partial \lambda} = \frac{n}{\lambda} - \sum t_i{}^{\alpha} = 0$

$$\Rightarrow \hat{\lambda} = \frac{n}{\sum t_i{}^{\hat{\alpha}}} = \frac{n}{\sum e^{\hat{\alpha} y_i}} \tag{3.2}$$

(3.1) and (3.2) are normal equations.

Substituting (3.2) in (3.1) and letting

$$g(\alpha) = \frac{1}{\alpha} + \frac{\sum y_i}{n} - \frac{\sum y_i e^{\alpha y_i}}{\sum e^{\alpha y_i}}, \tag{3.3}$$

Is evaluated $g(\alpha)$ at different values of $\alpha = 1.9$, 2.0, 2.025, 2.03515, 2.03525, 2.04, 2.05 and observing the change of value which will be MLE of α where $g(\alpha) = 0$ and then substituting $\hat{\alpha}$ in (3.2), we get $\hat{\lambda}$.

Also

$$-\left.\frac{\partial^2 logL}{\partial \alpha^2}\right|_{\hat{\alpha},\hat{\lambda}} = \frac{n}{\hat{\alpha}^2} + \hat{\lambda}\sum y_i{}^2 e^{\hat{\alpha} y_i}$$

$$-\left.\frac{\partial^2 logL}{\partial \alpha \partial \lambda}\right|_{\hat{\alpha},\hat{\lambda}} = -\sum y_i e^{\hat{\alpha} y_i}$$

$$-\left.\frac{\partial^2 logL}{\partial \lambda^2}\right|_{\hat{\alpha},\hat{\lambda}} = \frac{n}{\hat{\lambda}^2}$$

and so the estimate of variance-covariance matrix is

$$\begin{bmatrix} -\left.\frac{\partial^2 logL}{\partial \alpha^2}\right|_{\hat{\alpha},\hat{\lambda}} & -\left.\frac{\partial^2 logL}{\partial \alpha \partial \lambda}\right|_{\hat{\alpha},\hat{\lambda}} \\ -\left.\frac{\partial^2 logL}{\partial \alpha \partial \lambda}\right|_{\hat{\alpha},\hat{\lambda}} & -\left.\frac{\partial^2 logL}{\partial \lambda^2}\right|_{\hat{\alpha},\hat{\lambda}} \end{bmatrix} = \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{bmatrix}, \text{ say}$$

So that

$$\text{Var}(\hat{\alpha}) = \sigma_{11} \text{ and } \widehat{Var}(\hat{\alpha}) = \frac{\sigma_{11}}{n}, \widehat{s.e}(\hat{\alpha}) = \sqrt{\widehat{Var}(\hat{\alpha})}.$$

$$\text{Var}(\hat{\lambda}) = \sigma_{22} \text{ and } \widehat{Var}(\hat{\lambda}) = \frac{\sigma_{22}}{n}, \widehat{s.e}(\hat{\lambda}) = \sqrt{\widehat{Var}(\hat{\lambda})}.$$

$$\rho(\hat{\alpha},\hat{\lambda}) = \frac{\sigma_{12}}{\sqrt{\sigma_{11}\sigma_{22}}}$$

With the above simulated values from weibull distribution, we get following estimates ofparameter *a* and $\boldsymbol{\lambda}$ **and** their standard errors in Table 3.1

TABLE 2. MLE FOR SIMULATION STUDY OF WEIBULL DISTRIBUTION

Initial value of		$\hat{\alpha}$	$\hat{\lambda}$	Var($\hat{\alpha}$)	Var($\hat{\lambda}$)	S.E.($\hat{\alpha}$)	S.E.($\hat{\lambda}$)	Converge	-ll value	no. of iterations
α	λ									
1.90000	4.2	1.928085	4.244186	0.0004558805	0.005924449	0.02135136	0.07697044	0	636.1577	40
2.00000	4.2	2.001498	4.258776	0.0004891948	0.005970435	0.02211775	0.07726859	0	639.2733	41
2.02500	4.2	2.036173	4.261128	0.0005065795	0.005985999	0.02250732	0.07736924	0	637.3663	47
2.03515	4.2	2.076779	4.271012	0.0005287263	0.006030680	0.02299405	0.07765745	0	645.1097	38
2.03525	4.2	2.024185	4.138213	0.0005004243	0.005513361	0.02237017	0.07425201	0	569.3066	39
2.04000	4.2	2.018238	4.224875	0.0004952322	0.005817862	0.02225381	0.07627491	0	626.0774	40
2.05000	4.2	2.049999	4.264648	0.0005109614	0.005969990	0.02260445	0.07726571	0	649.7723	37

From table 3.1, we observe that the value of $\hat{\alpha}$ and $\hat{\lambda}$ are decreasing from $\hat{\alpha}$= 2.03525 and $\hat{\lambda}$= 4.138213 with their standard error as 0.02237017 and 0.07425201 respectively. Thus the simulation of Weibull distribution for 5000 values have given the above maximum likelihood estimates.

3.1 MLE OF COX MODEL WITH BASE LINE WEIBULL DISTRIBUTION AND GAMMA FRAILTY

Let survival time follows weibull distribution with parameters (γ,λ) with p.d.f. as

$$f(t) = \lambda\gamma t^{\gamma-1}e^{-\lambda t^{\gamma}} t > 0, \gamma > 0$$
$$= \gamma t^{\gamma-1}e^{-t^{\gamma}}, \text{ taking } \lambda=1.$$

and $\gamma \sim G(\alpha,1)$distribution as frailty distribution with p.d.f.

$$f(\gamma) = \frac{\gamma^{\alpha-1}e^{-\gamma}}{\Gamma\alpha}, \ \ \alpha > 0, \gamma > 0$$

Let $y_i = \log t_i$. Then the cox model reduces to

$$y_i = \beta_1 x_{i1} + \beta_2 x_{i2} + \gamma$$

so that $\gamma = y_i - \beta_1 x_{i1} - \beta_2 x_{i2}$

and likelihood will be

$$L = \prod_{i=1}^{n}(y_i - \beta_1 x_{i1} - \beta_2 x_{i2})^{\alpha-1} e^{-\sum(y_i-\beta_1 x_{i1}-\beta_2 x_{i2})}$$
$$logL = -nlog\Gamma\alpha + (\alpha-1)\sum log(y_i - \beta_1 x_{i1} - \beta_2 x_{i2}) - \sum(y_i - \beta_1 x_{i1} - \beta_2 x_{i2})$$

The normal equations are

$$\frac{\partial logL}{\partial\alpha} = \frac{-n\partial log\Gamma\alpha}{\partial\alpha} + \sum log(y_i - \beta_1 x_{i1} - \beta_2 x_{i2}) = 0 \quad (3.4)$$
$$\frac{\partial logL}{\partial\beta_1} = (\alpha-1)\sum\left(\frac{-x_{i1}}{y_i-\beta_1 x_{i1}-\beta_2 x_{i2}}\right) + \sum x_{i1} = 0 \quad (3.5)$$

$$\frac{\partial logL}{\partial \beta_2} = (\alpha - 1)\sum\left(\frac{-x_{i2}}{y_i - \beta_1 x_{i1} - \beta_2 x_{i2}}\right) + \sum x_{i2} = 0 \tag{3.6}$$

Solving above (3.4), (3.5) and (3.6) simultaneously with the kidney infection data given by McGilchrist and Aisbett (1991) of 38 kidney Patients with $Y = lnT$, the estimates of β_1(Age), β_2(Sex) and α with their standard errors are as under.

TABLE 3. THE ESTIMATES OF β_1(AGE), β_2(SEX) AND α WITH THEIR STANDARD ERROR

Parameter	Age	Sex	α
M.L.E	-0.0104088	1.4562985	3.3949261
S.E.	0.0117882	0.5256521	0.5129501

The estimate of σ^2 is 0.29455722, which is inverse of shape parameter of frailty distribution.

4. GENERALIZED EXPONENTIAL DISTRIBUTION AS FRAILTY MODEL

The Generalized exponential distribution is used in place of exponential distribution as frailty model. We give below the p.d.f., $f(t)$; survival function, $S(t)$; hazard function, $h(t)$ and cumulative hazard function, $H(t)$.

$$f(t) = \alpha\lambda(1 - e^{-\lambda t})^{\alpha-1}e^{-\lambda t}, \quad t > 0, \alpha > 0, \lambda > 0 \tag{4.1}$$

$$S_0(t) = \begin{cases} 1 - (1 - e^{-\lambda t})^{\alpha} & : t > 0, \alpha > 0, \lambda > 0 \\ 1 & : \quad o.w. \end{cases} \tag{4.2}$$

$$h_0(t) = \begin{cases} \frac{\alpha\lambda(1-e^{-\lambda t})^{\alpha-1}e^{-\lambda t}}{1-(1-e^{-\lambda t})^{\alpha}} & : t > 0, \alpha > 0, \lambda > 0 \\ 0 & : \quad o.w. \end{cases} \tag{4.3}$$

$$H_0(t) = \begin{cases} -ln[1 - \left(1 - e^{-\lambda t}\right)^{\alpha}] & : t > 0, \alpha > 0, \lambda > 0 \\ 0 & : \quad o.w. \end{cases} \tag{4.4}$$

4.1 SIMULATION AND ESTIMATION OF THE DISTRIBUTION

Let r $(0 < r < 1)$ be assumed. Since survival time $S(t)$ will be also in (0, 1) then equating $S(t) = r$ and since $S(t)$ for generalized exponential distribution is

$$S(t) = 1 - (1 - e^{-\lambda t})^{\alpha}$$
$$\Rightarrow 1 - (1 - e^{-\lambda t})^{\alpha} = r$$
$$\Rightarrow (1 - e^{-\lambda t})^{\alpha} = 1 - r$$
$$\Rightarrow log(1 - e^{-\lambda t}) = log(1 - r)^{\frac{1}{\alpha}}$$
$$\Rightarrow (1 - e^{-\lambda t}) = (1 - r)^{\frac{1}{\alpha}}$$
$$\Rightarrow e^{-\lambda t} = 1 - (1 - r)^{\frac{1}{\alpha}}$$
$$\Rightarrow -\lambda t = log\left[1 - (1 - r)^{\frac{1}{\alpha}}\right]$$
$$\Rightarrow t = -\frac{1}{\lambda} log\left[1 - (1 - r)^{\frac{1}{\alpha}}\right]$$
$$\Rightarrow t = log\left[1 - (1 - r)^{\frac{1}{\alpha}}\right]^{-\frac{1}{\lambda}}$$
$$= \frac{-log\left[1 - (1 - r)^{\frac{1}{\alpha}}\right]}{\lambda}$$

Simulating the values of t by giving values of $\alpha = 2$, $\lambda = 0.1$

$$S(t) = P(T > t) = 1 - F(t)$$
$$f(t) = F'(t)$$
$$L = \prod f(t_i)$$
$$logL = \sum logf(t_i)$$
$$= \sum logF'(t_i)$$
$$logL = -\sum logS'(t_i) \tag{4.5}$$

Since,

$$S(t) = 1 - (1 - e^{-\lambda t})^{\alpha}$$

Now,

$$\frac{\partial S(t)}{\partial t} = -\alpha(1-e^{-\lambda t})^{\alpha-1}\left(\lambda e^{-\lambda t}\right)$$

Using (4.5), we have

$$\begin{aligned} logL &= \sum log\left[\alpha(1-e^{-\lambda t_i})^{\alpha-1}\left(\lambda e^{-\lambda t_i}\right)\right] \\ &= \sum log\alpha + (\alpha-1)\sum log\left(1-e^{-\lambda t_i}\right) + \sum log\,\lambda + \sum loge^{-\lambda t_i} \\ &= nlog\alpha + (\alpha-1)\sum log\left(1-e^{-\lambda t_i}\right) + nlog\lambda + \sum loge^{-\lambda t_i} \\ &= nlog\alpha\lambda + (\alpha-1)\sum log\left(1-e^{-\lambda t_i}\right) - \lambda\sum t_i \end{aligned}$$

TABLE 4. MLE FOR SIMULATION STUDY OF GENERALIZED EXPONENTIAL DISTRIBUTION

Sample size = 50;	α	λ	σ^2
Estimates	1.5173151	0.0702266	0.659058
Standard error	0.3083377	0.0118815	
Sample size = 500;	α	λ	σ^2
Estimates	2.0404756	0.104021	0.490081
Standard error	0.1375155	0.0051668	
Sample size = 5000;	α	λ	σ^2
Estimates	1.9508322	0.0993446	0.512601
Standard error	0.0410658	0.0015708	

We observed that the estimates of α and λ are more stabilise when sample size increases. Thus $\hat{\alpha}$= 1.9508322 and $\hat{\lambda}$= 0.0993446 with their standard error as 0.0410658 and 0.0015708 respectively and σ^2= 0.512601.

5. EXPONENTIAL POWER DISTRIBUTION AS FRAITY

We consider exponential power distribution as more general form of exponential distribution and give below the p.d.f., $f(t)$; survival function, $S(t)$; hazard function, $h(t)$ and cumulative hazard function, $H(t)$.

$$f(t) = \alpha\lambda t^{\alpha-1}exp(\lambda t^{\alpha} + 1 - e^{\lambda t^{\alpha}}) \tag{5.1}$$

$$S_0(t) = \begin{cases} e^{\left[1-e^{\lambda t^{\alpha}}\right]} & : t > 0, \alpha > 0, \lambda > 0 \\ 1 & : \quad o.w. \end{cases} \tag{5.2}$$

$$h_0(t) = \begin{cases} \alpha\lambda t^{\alpha-1}e^{\lambda t^{\alpha}} & : t > 0, \alpha > 0, \lambda > 0 \\ 0 & : \quad o.w. \end{cases} \tag{5.3}$$

$$H_0(t) = \begin{cases} e^{\lambda t^{\alpha}} - 1 & : t > 0, \alpha > 0, \lambda > 0 \\ 0 & : \quad o.w. \end{cases} \tag{5.4}$$

5.1 SIMULATION AND ESTIMATION OF THE DISTRIBUTION

Let r $(0 < r < 1)$ be assumed. Since survival time $S(t)$ will be also in (0, 1) and hence $S(t)$ for Exponential power distribution by equating $S(t) = r$ is

$$S(t) = e^{1-e^{\lambda t^{\alpha}}}$$

$$\begin{aligned} e^{1-e^{\lambda t^{\alpha}}} = r &\Rightarrow 1 - e^{\lambda t^{\alpha}} = \log r \\ &\Rightarrow 1 - \log r = e^{\lambda t^{\alpha}} \\ &\Rightarrow \log(1-\log r) = \lambda t^{\alpha} \\ &\Rightarrow t = \left[\frac{\log(1-\log r)}{\lambda}\right]^{\frac{1}{\alpha}} \end{aligned}$$

We Generate 5000 values of t by using random numbers $r_1, r_2, \ldots, r_{5000} (0 < r < 1)$

For getting M.L.E., the likelihood will be derived as under

As $S(t) = e^{1-e^{\lambda t^{\alpha}}}$

$$\frac{\partial S(t_i)}{\partial t} = -\left(\alpha\lambda t_i{}^{\alpha-1} e^{\lambda t_i{}^{\alpha}+1-e^{\lambda t_i{}^{\alpha}}}\right)$$

$$log\left(-\frac{\partial S(t_i)}{\partial t}\right) = \log(\alpha\lambda) + (\alpha-1)logt_i + \lambda t_i{}^{\alpha} + 1 - e^{\lambda t_i{}^{\alpha}}$$

$$\sum \log\left(-\frac{\partial S(t_i)}{\partial t}\right) = \sum \log(\alpha\lambda) + (\alpha-1)\sum logt_i + \lambda \sum t_i{}^{\alpha} - \sum e^{\lambda t_i{}^{\alpha}} + n$$

$$logL = \sum \log\left(-\frac{\partial S(t_i)}{\partial t}\right) = n + n\log(\alpha\lambda) + (\alpha-1)\sum logt_i + \lambda \sum t_i{}^{\alpha} - \sum e^{\lambda t_i{}^{\alpha}}$$

$$logL = \sum \log\left(-\frac{\partial S(t_i)}{\partial t}\right) = n(1 + \log\alpha\lambda) + (\alpha-1)\sum logt_i + \lambda \sum t_i{}^{\alpha} - \sum e^{\lambda t_i{}^{\alpha}}$$

TABLE 5. MLE FOR SIMULATION STUDY OF EXPONENTIAL POWER DISTRIBUTION

Sample size = 50;	α	λ	σ^2
Estimates	1.8000427	0.0136814	0.555542
Standard error	0.2079574	0.0071355	
Sample size = 500;	α	λ	σ^2
Estimates	1.9544816	0.0113211	0.511644
Standard error	0.0726638	0.0019738	
Sample size = 5000;	α	λ	σ^2
Estimates	1.9655672	0.0108654	0.508758
Standard error	0.0235542	6.1228649×10^{-4}	

We observed that the estimates of α and λ are more stabilise when sample size increases. Thus,$\hat{\alpha}$=1.9655672and $\hat{\lambda}$=0.0108654with their standard error as 0.0235542and 6.1228649×10^{-4} respectively and σ^2= 0.508758

6. LOG-NORMAL FRAILTY MODELS

Log-normal frailty models are mostly used in modelling dependence structures in multivariate frailty models, e.g., in McGilchrist and Aisbett (1991), McGilchrist (1993), Lillard (1993), Lillard et al. (1995), Xue and Brookmeyer (1996), Sastry (1997), Gustafson (1997), Ripatti and Palmgren (2000); Ripatti et al. (2002), Huang and Wolfe (2002).

Flinn and Heckman (1982) have also applied the log-normal distribution in univariate cases. Two variants of the log-normal model exist. We assume that a normally distributed random variable W to generate frailty as $Z = e^{W}$. The two variants of the model are given by the restrictions **E**W = 0 and **E**Z = 1, where the first one is much more popular in the literature. Unfortunately, no explicit form of the unconditional likelihood exists. Consequently, estimation strategies based on numerical integration in the maximum likelihood approach are required.

6.1 SIMULATION AND ESTIMATION OF LOG NORMAL DISTRIBUTION

The p.d.f. of log normal distribution of t is

$$f(t;\, \mu, \sigma^2) = \frac{1}{\sqrt{2\pi}\,\sigma t} e^{-\frac{1}{2\sigma^2}(\log t-\mu)^2}$$

$$F(t) = \int_{-\infty}^{t} \frac{1}{\sqrt{2\pi}\,\sigma u} e^{-\frac{1}{2\sigma^2}(\log u-\mu)^2}$$

Let $\log u = y$

$$= \int_{0}^{logt} \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{1}{2\sigma^2}(y-\mu)^2 . y}$$

Let $w = \frac{y-\mu}{\sigma}$

$$F(t) = \int_{-\infty}^{\frac{logt-\mu}{\sigma}} \frac{1}{\sqrt{2\pi}} e^{-\frac{w^2}{2}} dw$$

Let $0 < r < 1$,

$$r = \int_{-\infty}^{\frac{logt-\mu}{\sigma}} \frac{1}{\sqrt{2\pi}} e^{-\frac{w^2}{2}} dw$$

6.2 ESTIMATION OF THE PARAMETERS OF LOGNORMAL DISTRIBUTION

Let $m_1' = e^{\mu+\frac{\sigma^2}{2}}$

$$var(t) = m_2 = e^{2\mu+\sigma^2}\left(e^{\sigma^2} - 1\right)$$
$$m_2 = m_1'^2\left(e^{\sigma^2} - 1\right)$$
$$\frac{m_2}{m_1'^2} = \left(e^{\sigma^2} - 1\right)$$
$$log\left(\frac{m_2}{m_1'^2} + 1\right) = \hat{\sigma}^2$$
$$logm_1' = \mu + \frac{\sigma^2}{2}$$
$$logm_1' - \frac{\hat{\sigma}^2}{2} = \hat{\mu}$$

Let $t_1, t_2, \ldots, t_{500}$ be generated values of log normal distribution and let $\bar{t} = \frac{\sum t_i}{n}$,
$m_2 = var(t) = \frac{1}{n}\sum\left(t_i - \bar{t}\right)^2$ then the moment estimators of μ and σ^2 are

$$\hat{\sigma}^2 = log\left(\frac{m_2}{\bar{t}^2}\right)$$
$$\hat{\mu} = log\bar{t} - \frac{\hat{\sigma}^2}{2}$$

6.3 ESTIMATION OF PARAMETERS OF LOG NORMAL DISTRIBUTION BY M.L.E.

The m.l.e. of μ and σ can be easily available as

$$\hat{\mu} = \frac{1}{n}\sum logt_i = \bar{y}$$

where $y_i = logt_i$, I = 1,2,…,500

$$\hat{\sigma}^2 = \frac{1}{n}\sum(y_i - \bar{y})^2$$

Now, we obtain the estimates of the parameters of μ, σ and their standard errors using simulation in the following table

TABLE 6. ESTIMATES OF THE PARAMETERS OF μ, σ AND THEIR STANDARD ERRORS

Initial value of μ	Initial value of σ	$\hat{\mu}$	$\hat{\sigma}$	Var($\hat{\mu}$)	Var($\hat{\sigma}$)	S.E.($\hat{\mu}$)	S.E.($\hat{\sigma}$)	Converge	ll value	no. of iterations
3	0.6	2.955624	0.527598	0.002783	0.001391	0.052759	0.037306	0	−281.620241	19

7. CONCLUSION

By considering different univariate continuous frailty models and using Monte Carlo method for simulated data, we obtained m.l.e. of the parameters of frailty models with their standard errors. Further the estimate of the variance of the base line distribution has been obtained.

REFERENCES

[1] Aalen, O.O. (1987). Two examples of modelling heterogeneity in survival analysis. Scandinavian Journal of Statistics 14, 19 – 25

[2] Abbring, J., van den Berg, G.J. (2005). The unobserved heterogeneity distribution in duration analysis. Working Paper, Free University Amsterdam

[3] Congdon, P. (1995). Modelling frailty in area mortality. Statistics in Medicine 14, 1859-74

[4] Flinn, C., Heckman, J. (1982). New methods for analyzing structural models of labor force dynamics. Journal of Econometrics 18, 115 – 168.

[5] Gustafson, P. (1997). Large hierarchical Bayesian analysis of multivariate survival data. Biometrics 53, 230 – 242.

[6] Hougaard, P. (2000). *Analysis of multivariate survival data*. Springer, New York.

[7] Huang, X., Wolfe, R.A. (2002). A frailty model for informative censoring. Biometrics 58, 510 – 520.

[8] Jeong, J.-H., Jung, S.-H., Wieand, S. (2003). A parametric model for long-term follow-up data from phase III breast cancer clinical trials. Statistics in Medicine 22, 339 – 352

[9] Jones, B.L. (1998). A model for analyzing the impact of selective lapsation on mortality. North American Actuarial Journal 2, 79 – 86

[10] Lancaster, T. (1979). Econometric methods for the duration of unemployment. Econometrica 47, 939 – 956

[11] Lillard, L.A. (1993). Simultaneous equations for hazards: marriage duration and fertility timing. Journal of Econometrics 56, 189 - 217

[12] Lillard, L.A., Brian, M.J., Waite, M.J. (1995). Premarital cohabitation and subsequent marital dissolution: a matter of self-selection? Demography 32, 437 – 457

[13] Manton, K.G., Stallard, E. (1981). Methods for evaluating the heterogeneity of aging processes in human populations using vital statistics data: Explaining the black/white mortality crossover by a model of mortality selection. Human Biology 53, 47 – 67

[14] Manton, K.G., Stallard, E., Vaupel, J.W. (1981). Methods for comparing the mortality experience of heterogeneous populations. Demography 18, 389 – 410.

[15] Manton, K., Stallard, E., Vaupel, J. (1986). Alternative models for heterogeneity of mortality risks among the aged. Journal of the American Statistical Association 81, 635 - 644

[16] McGilchrist, C.A., Aisbett, C.W. (1991). Regression with frailty in survival analysis. Biometrics 47, 461 – 466

[17] McGilchrist, C.A. (1993). REML estimation for survival models with frailty. Biometrics 49, 221-225

[18] Parekh S.G., Patel S.R., Ghosh D.K., and Raykundaliya D.P. (2016). Discrete Frailty Models. Proceedings of the 10th INDIACom; INDIACom-2016; IEEE conference ID: 37465. 2016 3rd International Conference on "Computing for Sustainable Global Development", 180-182.

[19] Ripatti, S., Palmgren, J. (2000). Estimation of multivariate frailty models using penalised partial likelihood. Biometrics 56, 1016 - 1022

[20] Ripatti, S., Larsen, K., Palmgren, J. (2002). Maximum likelihood inference for multivariate frailty models using an automated MCEM algorithm. Lifetime Data Analysis 8, 349 - 360

[21] Sastry, N. (1997). A nested frailty model for survival data, with an application to the study of child survival in northeast Brazil. Journal of the American Statistical Association 92, 426 – 435

[22] Vaupel, J.W., Manton, K.G., Stallard, E. (1979). The impact of heterogeneity in individual frailty on the dynamics of mortality. Demography 16(3): 439 – 454

[23] Xue, X., Brookmeyer, R. (1996). Bivariate frailty model for the analysis of multivariate survival time. Lifetime Data Analysis 2, 277 – 290