

# We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900

Open access books available

186,000

International authors and editors

200M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index  
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?  
Contact [book.department@intechopen.com](mailto:book.department@intechopen.com)

Numbers displayed above are based on latest data collected.  
For more information visit [www.intechopen.com](http://www.intechopen.com)



# Modeling Heterogeneity Using Lindley Distribution

*Arvind Pandey and Lalpawimawha*

## Abstract

Frailty models are intended for use in survival analysis to explain unobserved heterogeneity in an individual caused by various hereditary variables or environmental influences. A shared frailty model was utilized to examine the data. It is based on the idea that frailty affects the hazard rate in a multiplicative manner. In this manuscript, we introduce a new frailty model called the Lindley shared frailty model with exponential power and generalized Rayleigh as baseline distributions. The Bayesian method of the Monte Carlo method of the Markov chain is used to estimate the parameters used in the model; simulation studies are also carried out to compare the actual and calculated values of the parameters; the proposed model is compared with the Bayesian comparison method. Compare and propose the best model of infectious disease data.

**Keywords:** Bayesian technique, exponential power distribution, generalized Rayleigh distribution, Lindley frailty, MCMC

## 1. Introduction

The term frailty was coined by Vaupel et al. [1]. The frailty model is typically represented as an unobservable random variable that multiplies the risk function, with the frailty random variable supposed to be one of the parameter distributions, such as gamma, log-normal, positive stable, inverse Gaussian, power variance function, and so on. Let  $Y$  be a continuous random variable of lifetime of an individual and the frailty random variable (RV) be  $V$ . The conditional hazard function (CHF) for a given frailty variable  $V = v$  at time  $y > 0$  is

$$m(y|v) = v h_0(y) e^{X'\beta}, \quad (1)$$

where  $m_0(y)$  is a baseline hazard function (BHF) at time  $y > 0$ ,  $X$  is a covariate and  $\beta$  is a regression coefficient, these are in vector form. The CHF for given frailty at time  $y > 0$  is

$$S(y|v) = e^{-\int_0^y m(x|v) dx} = e^{-v M_0(y) e^{X'\beta}}, \quad (2)$$

where  $M_0(y)$  is cumulative baseline hazard function (CBHF) at time  $y > 0$ . Integrating over the range of frailty variable  $V$  having density  $f(v)$ , we get marginal survival function as

$$S(y) = \int_0^{\infty} S(y|v)f(v)dv, \quad (3)$$

$$= L_v\left(M_0(y)e^{X'\beta}\right), \quad (4)$$

where  $L_V(\cdot)$  is a Laplace transformation of the distribution of  $V$ . Once we have survival function at time  $y > 0$  of lifetime random variable of an individual one can obtain probability structure and can base their inference on it.

Frailty models have gained more attention in the recent medical research due to the uniqueness property of the frailty parameter. Generally, gamma distribution, log-normal distribution and inverse Gaussian distributions are the most commonly used frailty distributions [2, 3]. Hanagal and Dabade [4] introduced new Compound negative binomial shared frailty models for bivariate survival data using Weibull and generalized exponential as baseline distributions. Pandey et al. [5] compared gamma, inverse Gaussian and positive frailty models with generalized Pareto as baseline distribution. Pandey et al. [6] also compared gamma and inverse Gaussian frailty distributions under additive property.

To extract the features of the Lindley shared frailty model, we used Lindley as a frailty distribution with right censored data under generalized Rayleigh and exponential power as baseline distributions. The survival periods are dependent in this case because the frailty variable follows the Lindley distribution. The predicted value of the frailty distribution variance influences the population's degree of heterogeneity. The higher the variance of the frailty distribution, the more heterogeneity there is in the population under consideration. The frailty distribution becomes degraded when zero variance is observed. As a baseline distribution, the exponential power distribution is used. Because it exhibits a rising hazard rate, which is typical in real-life distributions, the exponential power distribution is chosen as the baseline distribution. The Lindley distribution with one parameter was first proposed by Lindley [7] for analyzing failure times data. It belongs to an exponential family, but it is used as an alternative to the exponential distribution. Lindley distribution is alluring due to the ability of modeling failure time data with increasing, decreasing, unimodal and bathtub shaped hazard rates. Ghitany et al. [8, 9] discussed different properties of Lindley distribution and also showed that Lindley distribution is better than the exponential distribution for modeling failure time data when considering hazard rate is unimodal or bathtub shaped. It is also shown that Lindley distribution is more flexible than exponential distribution in modeling lifetime data. Many authors have discussed and introduced different generalization of Lindley distribution. Bakouch et al. [10] introduced extended Lindley distribution. Ghitany et al. [11] proposed the power Lindley distribution and Shanker et al. [12] proposed two parameter Lindley distribution, which could also be reduced to one parameter case. The mean of a two parameter Lindley distribution is always greater than the mode indicating that the distribution is positively skewed.

The classic approach and the Bayesian approach are two widely utilized techniques in general. We can employ prior distributions here, therefore we'll estimate the model parameters using the Bayesian Markov Chain Monte Carlo (MCMC) approach. Furthermore, because characteristics with diverse posterior distributions may be easily generated, the results and model selection criteria can be clearly interpreted. Run after thinning mean and autocorrelation plots, follow-up plots, past plot couplings, sample autocorrelation plots dictate chain behavior, burn duration, autocorrelation delay, and how observations are made It's utilized for cognitive confirmation on its own. We also give simulation experiments to back up the model's performance. All of the model's estimation processes are detailed, as well as infection statistics relating to kidney infections.

In Sections 2 and 3, the introduction of the Lindley shared frailty model and baseline distributions are given, followed by proposed models and estimation strategies in Sections 4 and 5. In Sections 6 and 7, application of the proposed model and discussion are given.

## 2. Lindley shared frailty model

Let a continuous random variable  $V$  follows two parameter Lindley distribution (TPLDP) with parameters  $\alpha$  and  $\lambda$  then density function of  $V$  is

$$f(v) = \begin{cases} \frac{\alpha^2}{\alpha\lambda + 1} (\lambda + v)e^{-\lambda v} & ; v > 0, \alpha > 0, \lambda\alpha > -1 \\ 0 & ; otherwise, \end{cases} \quad (5)$$

and the Laplace transform is

$$L_V(s) = \frac{\alpha^2(1 + (s + \alpha)\lambda)}{(s + \alpha)^2(1 + \lambda\alpha)}, s + \alpha > 0. \quad (6)$$

The mean and variance of frailty variable are  $E(Z) = \frac{\alpha\lambda+2}{\alpha(\alpha\lambda+1)}$  and  $V(V) = \frac{2+4\alpha\lambda+\alpha^2\lambda^2}{\alpha^2(\alpha\lambda+1)^2}$ . For identifiability, we assume  $V$  has expected value equal to one i.e.  $E(V) = 1$ , which imply that  $\alpha = \xi$  and  $\lambda = \frac{2-\xi}{\xi(\xi-1)}$ . Under this restriction the density function and the Laplace transformation of Lindley distribution reduces to

$$f(v) = \begin{cases} \frac{e^{-\xi v} \xi (\xi^2 v + \xi - \xi v - 2)}{\xi - 2} & ; v > 0, \xi > 0 \\ 0 & ; otherwise, \end{cases} \quad (7)$$

and the Laplace transform is

$$L_V(s) = \frac{\xi(\xi + s(2 - \xi))}{(s + \xi)^2}. \quad (8)$$

with variance of  $V$  is  $\frac{4\xi - \xi^2 - 2}{\xi^2}$ . The frailty variable  $V$  is degenerate at  $V = 1$ . Replacing Laplace transform in Eq. (4), we get the unconditional bivariate survival function for the  $j^{th}$  individual as

$$S(y_{1k}, y_{2k}) = \frac{\xi(\xi + \eta_k(M_{01}(y_{1k}) + M_{02}(y_{2k}))(2 - \xi))}{(\eta_k(M_{01}(y_{1k}) + M_{02}(y_{2k})) + \xi)^2} \quad (9)$$

where  $M_{01}(y_{1k})$  and  $M_{02}(y_{2k})$  are the cumulative baseline hazard functions of the lifetime  $Y_{1k}$  and  $Y_{2k}$  respectively.

And for without frailty, the model becomes

$$S(y_{1k}, y_{2k}) = e^{-\eta_k(M_{01}(y_{1k}) + M_{02}(y_{2k}))}. \quad (10)$$

### 3. Baseline distributions

As a starting point, we'll look at the generalized Rayleigh distribution. Surles and Padgett [13] proposed the two-parameter Burr type X distribution, dubbed the generalized Rayleigh distribution, and demonstrated that the two-parameter generalized Rayleigh distribution may be utilized to describe strength and general lifetime data rather efficiently. The two-parameter generalized Rayleigh distribution can be utilized well in survival analysis to describe strength data as well as general lifetime data. If a continuous random variable  $Y$  has a two-parameter generalized Rayleigh distribution, the survival function, hazard function, and cumulative hazard function are as follows:

$$S(y) = 1 - \left(1 - e^{-(\lambda y)^2}\right)^\alpha; y > 0, \lambda > 0, \alpha > 0 \quad (11)$$

$$m(y) = \frac{2\alpha\lambda^2 y e^{-(\lambda y)^2} \left(1 - e^{-(\lambda y)^2}\right)^{\alpha-1}}{1 - \left(1 - e^{-(\lambda y)^2}\right)^\alpha}; y > 0, \lambda > 0, \alpha > 0 \quad (12)$$

$$M(y) = -\log \left[1 - \left(1 - e^{-(\lambda y)^2}\right)^\alpha\right]; y > 0, \lambda > 0, \alpha > 0 \quad (13)$$

where  $\alpha$  and  $\lambda$  stands for shape and scale parameters respectively of the distribution. It has also some attractive properties increasing hazard and bathtub type depends on the parameter value.

The second baseline distribution considered here is exponential power distribution. A continuous random variable  $Y$  is said to follow the exponential power distribution if its survival function, hazard function and cumulative hazard function are, respectively,

$$S(y) = e^{1-e^{\lambda y^\alpha}}; y > 0, \lambda > 0, \alpha > 0 \quad (14)$$

$$m(y) = \alpha\lambda y^{\alpha-1} e^{\lambda y^\alpha}; y > 0, \lambda > 0, \alpha > 0 \quad (15)$$

$$M(y) = e^{\lambda y^\alpha} - 1 \quad (16)$$

where  $\lambda$  and  $\alpha$  are the shape and scale parameters of the exponential power distribution. The hazard function and cumulative hazard function are respectively,

$$m(y) = \alpha\lambda y^{\alpha-1} e^{\lambda y^\alpha}; y > 0, \lambda > 0, \alpha > 0 \quad (17)$$

$$M(y) = e^{\lambda y^\alpha} - 1 \quad (18)$$

The hazard function is decreasing function at time  $y$  when  $\alpha < 1$  for smaller values of  $\lambda$  but as  $\lambda$  increases hazard function takes  $U$  shape curve and further increment in  $\lambda$  gives increasing nature to hazard function.

### 4. Proposed models

The unconditional survival function is obtained by replacing the cumulative hazard functions of generalized Rayleigh distribution and exponential power distribution in Eqs. (9) and (10). Then,

$$S(y_{1k}, y_{2k}) = e^{-\left(\left(-\log \left[1 - \left(1 - e^{-(\lambda_1 y_{1k})^2}\right)^{\alpha_1}\right]\right) + \left(-\log \left[1 - \left(1 - e^{-(\lambda_2 y_{2k})^2}\right)^{\alpha_2}\right]\right)\right)} \eta_k \left[1 + \xi \left(\left(-\log \left[1 - \left(1 - e^{-(\lambda_1 y_{1k})^2}\right)^{\alpha_1}\right]\right) + \left(-\log \left[1 - \left(1 - e^{-(\lambda_2 y_{2k})^2}\right)^{\alpha_2}\right]\right)\right)\right]^{-1/\xi} \quad (19)$$

$$S(y_{1k}, y_{2k}) = e^{-\left(\frac{\alpha_1}{\lambda_1} (e^{\lambda_1 y_{1k}} - 1) + \frac{\alpha_2}{\lambda_2} (e^{\lambda_2 y_{2k}} - 1)\right)} \eta_k \left[1 + \xi \left(\frac{\alpha_1}{\lambda_1} (e^{\lambda_1 y_{1k}} - 1) + \frac{\alpha_2}{\lambda_2} (e^{\lambda_2 y_{2k}} - 1)\right)\right]^{-1/\xi} \quad (20)$$

$$S(y_{1k}, y_{2k}) = e^{-\left(\left(-\log \left[1 - \left(1 - e^{-(\lambda_1 y_{1k})^2}\right)^{\alpha_1}\right]\right) + \left(-\log \left[1 - \left(1 - e^{-(\lambda_2 y_{2k})^2}\right)^{\alpha_2}\right]\right)\right)} \eta_k \quad (21)$$

$$S(y_{1k}, y_{2k}) = e^{-\left(\frac{\alpha_1}{\lambda_1} (e^{\lambda_1 y_{1k}} - 1) + \frac{\alpha_2}{\lambda_2} (e^{\lambda_2 y_{2k}} - 1)\right)} \eta_k \quad (22)$$

The Eqs. (19) and (20) are Lindley shared frailty model with generalized Rayleigh and exponential power as baseline distributions, called as Model-I and Model-II and Eqs. (21) and (22) are without frailty models under the same baseline distributions, called as Model-III and Model-IV.

## 5. Estimation strategies

By assuming independence between censoring scheme and individual lifetimes, the likelihood function associated with failure times for the  $k^{th}$  people ( $k = 1, 2, 3, n$ ) and censoring times is given by

$$I(\underline{\Psi}, \beta, \xi) = \prod_{k=1}^{n_1} f_1(y_{1k}, y_{2k}) \prod_{k=1}^{n_2} f_2(y_{1k}, d_{2k}) \prod_{k=1}^{n_3} f_3(d_{1k}, y_{2k}) \prod_{k=1}^{n_4} f_4(d_{1k}, d_{2k}) \quad (23)$$

where  $\underline{\Psi}$ ,  $\beta$  and  $\xi$  are vectors of baseline parameters, regression coefficients and frailty distribution parameter. The likelihood function for without frailty is given as

$$I(\underline{\Psi}, \beta) = \prod_{k=1}^{n_1} f_1(y_{1k}, y_{2k}) \prod_{k=1}^{n_2} f_2(y_{1k}, d_{2k}) \prod_{k=1}^{n_3} f_3(d_{1k}, y_{2k}) \prod_{k=1}^{n_4} f_4(d_{1k}, d_{2k}) \quad (24)$$

and  $n_1, n_2, n_3$  and  $n_4$  are the number of observations, which are observed to lie in the intervals  $y_{1k} < d_{1k}, y_{2k} < d_{2k}; y_{1k} < d_{1k}, y_{2k} > d_{2k}; y_{1k} > d_{1k}, y_{2k} < d_{2k}$  and  $y_{1k} > d_{1k}, y_{2k} > d_{2k}$  respectively and the contribution of the  $k^{th}$  individual in the likelihood function as

$$\begin{aligned} f_1(y_{1k}, y_{2k}) &= \frac{\partial^2 S(y_{1k}, y_{2k})}{\partial y_{1k} \partial y_{2k}} \\ f_2(y_{1k}, d_{2k}) &= -\frac{\partial S(y_{1k}, d_{2k})}{\partial y_{1k}} \\ f_3(d_{1k}, y_{2k}) &= -\frac{\partial S(d_{1k}, y_{2k})}{\partial y_{2k}} \\ f_4(d_{1k}, d_{2k}) &= S(d_{1k}, d_{2k}) \end{aligned} \quad (25)$$

Putting Eq. (24) in Eqs. (23) and (24), we get the likelihood functions for the Lindley shared frailty models under generalized Rayleigh and exponential power baseline distributions and likelihood function for without frailty models under the same baseline distributions.

The joint posterior density of the parameters given failure times is given as

$$\pi(\alpha_1, \lambda_1, \alpha_2, \lambda_2, \xi, \beta) \propto L(\alpha_1, \lambda_1, \alpha_2, \lambda_2, \xi, \beta) \times g_1(\alpha_1)g_2(\lambda_1)g_3(\alpha_2)g_4(\lambda_2)g_5(\xi) \prod_{i=1}^5 p_i(\beta_i)$$

where  $g_i(\cdot)$  ( $i = 1, 2, \dots, 5$ ) represent the prior density function of baseline parameters and frailty variance, which are suppose to have known hyper parameters;  $p_i(\cdot)$  represents prior density function for the regression coefficient  $\beta_i$ ;  $\beta_i$  represents regression coefficients of vector form except  $\beta_i, i = 1, 2, \dots, a$  and likelihood function  $I(\cdot)$  is also presented by Eqs. (23) and (24). It is assumed that all of the parameters are distributed independently in this case.

The expression of the likelihood function in Eqs. (23) and (24) are not easy to solve by using the Newton–Raphson method. MLEs fail to converge as it involved a large number of parameters. As a result, the Bayesian approach was used to estimate the parameters involved in the models, which is free of such issues.

Prior distributions are utilized as follows: for a frailty parameter with a small value of  $\Psi$ , a gamma distribution with mean 1 and big variance  $\Gamma(\Psi, \Psi)$  is used as a prior distribution. As a prior for the regression coefficient, say  $\varphi^2$ , a normal distribution with mean zero and huge variance is utilized. Because we do not know anything about the baseline parameters, we use the same type of prior distributions used by Ibrahim et al. [14] and Sahu et al. [15], as well as a non-informative prior. As non-informative prior distributions,  $\Gamma(a_1, b_1)$  and  $U(a_2, b_2)$  are utilized. All the hyper-parameters  $\Psi, \varphi, a_1, a_2, b_1$  and  $b_2$  are supposed to be known in advanced. Here  $\Gamma(a_1, b_1)$  stands for gamma distribution having shape parameter  $a_1$  and scale parameter  $b_1$  and  $U(a_2, b_2)$  stands for the uniform distribution over the interval  $a_2$  to  $b_2$ . We provide the hyper-parameters as  $\Psi = 0.0001, \varphi^2 = 1000, a_1 = 1, b_1 = 0.0001, a_2 = 0$ , and  $b_2 = 100$ .

The Metropolis Hasting Algorithm and Gibbs Sampler were used to estimate the parameters in the models fitted with the preceding prior density function and likelihood Eqs. (23) and (24), Metropolis Hasting Algorithm and Gibbs Sampler was utilized. Geweke test and Gelman–Rubin statistics, as suggested by Geweke [16] and Gelman et al. [17], show that the Markov chain converges to a stationary distribution. We used trace plots, coupling from the past plots, and sample autocorrelation plots to examine the chain’s behavior, as well as to determine the burn-in period and autocorrelation lag.

It is important to decide which model provides the best fit to the dataset, the comparison of models was done using Akaike Information Criteria (AIC), Bayesian Information Criteria (BIC), Deviance Information Criteria (DIC) and Bayes factor.

## 6. Application in real life data

The models’ applicability was tested by applying them to infectious illness data relating to kidney infection that occurred during catheter implantation [18]. It includes 38 patients’ first and second recurrence times of infection from catheters used with portable dialysis equipment. For each patient in a cluster, these two times of infection are clustered together. Other pertinent data includes infection duration,

patient age, gender (0 for male and 1 for female), and illness kinds such as Glomerulo Nephritis (GN), Acute Nephritis (AN), and Polycystic Kidney Disease (PKD).

To begin, the Kolmogorov–Smirnov test is used to determine the goodness of fit for kidney infection data, and the p-values obtained for the first and second recurrences are large enough to rule out the hypothesis that the first and second recurrence times follow the distributions with survival functions as given in Eqs. (11) and (14) in univariate case and it is also assumed to be appropriate for bivariate case. The corresponding p-values are given in **Table 1**. The posterior summary of the proposed models are presented in **Tables 2** and **3**. It consists of estimate (posterior mean), standard error, 95% lower and upper credible limits, GR statistics values with p-values and Geweke test values. From **Tables 2** and **3**, It is observed that we can observe that regression coefficients for all the models are more or less same. Also for all these proposed models, the value zero is not a credible value for all the credible intervals of the regression coefficients, so all the covariates are seems to be significant. To test the models' accuracy, we created 95% and 50% predictive intervals from a generated random sample based on a predictive distribution as described by Sahu et al. [15], and counted the total number of actual recurrence times for first and second kidney infections that fell within the intervals. The 95 percent and 50 percent predictive intervals are contained in the 95% and 50% predictive intervals for Models I and II, respectively, 76, 58, and 76, 60 out of 76

Distribution	Recurrence times	
	First	Second
Generalized Rayleigh	0.98078	0.99889
Exponential power	0.96291	0.75766

**Table 1.**  
*p-Values of K-S Statistics for goodness of fit test for kidney infection data set.*

Parameter	Estimate	Standard error	Lower credible limit	Upper credible limit	Geweke values	p values	Gelman & Rubin values
<b>Burn in period = 5600; autocorrelation lag = 275</b>							
$\alpha_1$	0.3716	0.0312	0.3118	0.4283	1.0007	-0.0048	0.4980
$\alpha_2$	0.4253	0.0455	0.3326	0.5044	1.0003	-0.0045	0.4981
$\lambda_1$	0.0032	0.0004	0.0023	0.0041	1.0008	-0.0017	0.4992
$\lambda_2$	0.0026	0.0004	0.0018	0.0034	1.0031	-0.0052	0.4979
$\xi$	1.1287	0.0423	1.0722	1.2196	1.0032	-0.0095	0.4961
$\beta_1$	0.0153	0.0041	0.0083	0.0238	1.0052	-0.0042	0.4983
$\beta_2$	-1.0792	0.2740	-1.6013	-0.5350	1.0001	0.0070	0.4983
$\beta_3$	0.0021	0.0004	0.0012	0.0029	1.0008	-0.0060	0.4975
$\beta_4$	0.0031	0.0004	0.0022	0.0040	1.0005	-0.0010	0.4995
$\beta_5$	-0.2149	0.0514	-0.3031	-0.0947	1.0008	0.0059	0.5023

**Table 2.**  
*Posterior results with baseline generalized Rayleigh distribution.*

Parameter	Estimate	Standard error	Lower credible limit	Upper credible limit	Geweke values	p values	Gelman & Rubin values
<b>Burn in period = 6800; autocorrelation lag = 280</b>							
$\alpha_1$	0.4440	0.0251	0.3878	0.4905	1.0010	-0.0065	0.4973
$\alpha_2$	0.5040	0.0368	0.4263	0.5689	1.0002	0.0032	0.5012
$\lambda_1$	0.3120	0.0471	0.2353	0.4021	1.0001	0.0029	0.5011
$\lambda_2$	0.2150	0.0445	0.1530	0.3166	1.0016	-0.0053	0.4978
$\xi$	1.2061	0.0496	1.1197	1.3013	0.9999	0.0026	0.5010
$\beta_1$	0.0001	0.0001	1.7e-05	0.0002	1.0003	0.0036	0.5014
$\beta_2$	-2.5247	0.3867	-3.2854	-1.7454	1.0021	0.0071	0.5014
$\beta_3$	0.0020	0.0004	0.0012	0.0029	1.0006	0.0119	0.5047
$\beta_4$	0.0031	0.0004	0.0021	0.0040	1.0003	0.0107	0.5042
$\beta_5$	-0.9916	0.4466	-1.8481	-0.1704	1.0027	0.0001	0.5000

**Table 3.**  
Posterior results with baseline exponential power distribution.

observations. This demonstrates that the two models are appropriate for the data. Model-I is a better model in terms of AIC, BIC, and DIC values, since it has lower AIC, BIC, and DIC values than Model-II in **Table 4**. However, because the difference between AIC, BIC, and DIC values for Model I and Model II is so little, AIC, BIC, and DIC values are not suitable for deciding between the two models. To compare model  $u$  with model  $v$ , we use the Bayes factor (**Table 5**). Model-I is better than Model-II, since the equivalent value of  $2 \log(B_{uv})$  is larger than 10, suggesting that there is a very strong positive to favor Model-I over Model-II for the provided dataset, confirming our earlier conclusion in **Table 4**. As a result of all of the demonstrated comparison criteria, we can conclude that Model-I is superior to Model-II in terms of modeling kidney infection data.

Model no.	AIC	BIC	DIC
Model I	638.5262	654.9020	625.1190
Model II	700.3005	716.6763	686.8069
Model III	691.5817	706.3200	720.7843
Model IV	702.0827	716.8210	689.8978

**Table 4.**  
AIC, BIC and DIC values for all models.

Numerator model against denominator model	$2 \log_e(B_{uv})$	Range	Evidence against model in denominator
Model - I against Model - II	63.23936	> 10	Very Strong Positive

**Table 5.**  
Bayes factor values and decision for test of significance for frailty fitted to kidney infection data set.

## 7. Discussion

In this study, we examined a new Lindley shared frailty model under generalized Rayleigh and exponential power as baseline distributions.

To suit all of the proposed models, the Metropolis-Hastings and Gibbs sampler was used. The proposed models were used to assess kidney infection data, and the best model was suggested. To conduct the analysis, we used self-composed programs in the R statistical software.

All of the exhibited comparison criteria indicated that the Lindley shared frailty model with generalized Rayleigh baseline distribution is superior to exponential power baseline distribution and without frailty models for modeling kidney infection data under the identical baseline distributions. The estimates of frailty variance are 0.9415 and 0.9739, which are high in all the proposed models indicating that there is a strong evidence of a high degree of heterogeneity among the patients in the population. A few patients are anticipated to be exceptionally inclined to infection compared to others with the same covariate values. Some patients are expected to be very prone to infection compared to others with the same covariate values. Also we can say that there is a strong positive correlation between the two infection times for the same patient.

The most important properties of the proposed models that were not mentioned in the previous study are the estimates of the frailty variances are high in all proposed models as compared to previous study given by McGilchrist and Aisbett [18] on log-normal frailty, Hanagal and Bhambure [19], the disease type GN and AN has lower infection rates as compared to other covariates. All the covariates are significant factors for kidney infection, but the disease type are insignificant in the previous proposed frailty models (see [4]). It is very crucial to be mention that Lindley shared frailty model based on generalized Rayleigh baseline distribution is performed better to analyze kidney infection data than other frailty models [4, 19].

### Author details

Arvind Pandey<sup>1</sup> and Lalpawimawha<sup>2\*</sup>

<sup>1</sup> Department of Statistics, Central University of Rajasthan, Rajasthan, India

<sup>2</sup> Department of Statistics, Pachhunga University College, Mizoram, India

\*Address all correspondence to: [raltelalpawimawha08@gmail.com](mailto:raltelalpawimawha08@gmail.com)

### IntechOpen

© 2021 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. 

## References

- [1] 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.
- [2] Gupta, R.D., and Kundu, D. (2001). Exponentiated Exponential Family: An Alternative to Gamma and Weibull Distributions, *Biometrical Journal*, 43(1), 117-130.
- [3] Kheiri, S., A. Kimber, and M.R. Meshkani. 2007. Bayesian analysis of an inverse Gaussian correlated frailty model. *Computational Statistics and Data Analysis* 51: 5317-5326.
- [4] Hanagal, D.D., Dabade, A.D. 2013. Compound negative binomial shared frailty models for bivariate survival data. *Statistics and Probability Letters*, 83, 2507-2515.
- [5] Pandey, A., Bhushan, S., Lalpawimawha, R. 2018. Shared frailty models with baseline generalized Pareto distribution, *Communications in Statistics-Theory and Methods*, DOI: 10.1080/03610926.2018.1500597
- [6] Pandey, A., Bhushan, S., Lalpawimawha, R., Misra, P.K. 2019. Comparison of additive shared frailty models under Lindley baseline distribution, *Communications in Statistics-Simulation and Computation*, DOI: 10.1080/03610918.2019.1664573
- [7] Lindley, D.V. 1958. Fiducial distributions and Bayes' theorem. *Journal of the Royal Statistical Society, Series B*, 20, 102-107.
- [8] Ghitany, M.E., Atieh, B., Nadarajah, S. 2008. Lindley Distribution and Its Applications. *Mathematics and Computers in Simulation*, 78(4), 2008, 493-506.
- [9] Ghitany, M.E., Alqallaf, F., Al-Mutairi, D.K., Hussain, H. 2011. A Two Parameter Weighted Lindley Distribution and Its Applications to Survival Data. *Mathematics and Computers in Simulation*, 81(6), 1190-1201.
- [10] Bakouch, H.S., Al-Zahrani, B.M., Al-Sho, A.A., Marchi, A.A., Louzada, F. 2012. An Extended Lindley Distribution. *Journal of the Korean Statistical Society*, 41(1), 75-85.
- [11] Ghitany, M., Al-Mutairi, D., Balakrishnan, N. and Al-Enezi, I., 2013. Power Lindley distribution and associated inference. *Computational Statistics and Data Analysis*, 64, 20-33.
- [12] Shanker, R., Sharma, S., Shanker, R. 2013. A Two-Parameter Lindley Distribution for Modeling Waiting and Survival Times Data. *Applied Mathematics*, 4, 363-368.
- [13] Surlles, J. G., Padgett, W.J. (2001). Inference for reliability and stress-strength for a scaled Burr Type X distribution. *Lifetime Data Anal.* 7,187-200.
- [14] Ibrahim, J.G., Chen, Ming-Hui, Sinha, D., 2001. *Bayesian Survival Analysis*. Springer-Verlag.
- [15] Sahu, S.K., D.K. Dey, H. Aslanidou, & D. Sinha. 1997. A Weibull regression model with gamma frailties for multivariate survival data. *Life time data analysis*, 3, 123-137.
- [16] Geweke, J. 1992. "Evaluating the Accuracy of Sampling-Based Approaches to the Calculation of Posterior Moments." In *Bayesian Statistics 4* (eds. J.M. Bernardo, J. Berger, A.P. Dawid and A.F.M. Smith), 169-193. Oxford: Oxford University Press.

[17] Gelman, A., & D.B. Rubin. 1992. A single series from the Gibbs sampler provides a false sense of security. In *Bayesian Statistics 4* (J. M. Bernardo, J. O. Berger, A. P. Dawid and A. F. M. Smith, eds.). Oxford: Oxford Univ. Press, 625-632.

[18] McGilchrist, C.A., & C.W. Aisbett. 1991. Regression with frailty in survival analysis. *Biometrics*, 47, 461-466.

[19] Hanagal, D.D., & Bhambure, S.M. (2016). Modeling bivariate survival data using shared inverse Gaussian frailty model. *Communications in Statistics-Theory and Methods*, 45(17), 4969-4987.

IntechOpen