ISSN: 2278-4632
Vol-12 Issue-01 No.01: 2022

Shared Gamma Frailty Model in Survival Analysis

Raman T T
Professor, Department of Mathematics
St.Joseph's Institute of Technology, Chennai, India.

Abstract

Introduction:

Breast cancer is frightening the women around the globe and its causes have been studied worldwide. It is necessary for the people affected with the breast cancer to identify the risk factors and prognostic factors of breast cancer for their survival, and this research is specific to the south Indian population. This study attempts to explore the survival experience of various set of covariates with respect to stages of breast cancer patients using the Shared gamma frailty model.

Keywords: Breast Cancer, Shared Frailty, Frailty model, Gamma distribution.

Objective:

To identify the better survival model using shared gamma frailty model with respect to disease stage of different set of covariates.

Methodology:

The data obtained from Government Hospital includes 522 women, diagnosed and treated with adjuvant and neo-adjuvant therapy between January 2000 - December 2008 and follow up period up to May 2010 is used for comparison of different set of covariates with respect to Cancer stage of Share gamma frailty model.

Results:

The model-I includes the covariates chemotheraphy, menopause status, family history of cancer, this model shows chemotherapy is significant factor and deviance is 5184.In model-II includes the additional covariates number of children, number of abortions, and the deviance becomes less 5050, but it reflects the same significant covariates. In model- III includes one more additional covariates of age which doesn't contribute to the model, because the deviance value is increased to 5145.6.So model-II covariates yields better results for the breast cancer data.

Conclusion:

The shared gamma frailty model with respect to stages of breast cancer, model-II covariates accounts more heterogeneity compared with other models.

Shared Gamma Frailty Models

Frailty models have been used when groups of subjects have responses that are likely to be dependent in some general way. When multiple events have been observed on the same subjects (Liang et al., 1995) discuss the use of frailty models with multivariate failure time data. In this background, if the value of the frailty is assumed to be constant within groups, the models are called shared frailty models. The shared frailty model has been extended by (Picklets, et al., 1994; Yashin, et al., 1995) to allow difference but correlated frailties among observations within a group. The concept of frailty model has been proposed for use in generalized linear models by Clayton (1994) and Neunaus (1992).

In a shared frailty model, the conditional hazard function of T_{ij}, given the unobservable frailty random variable Y_i of the i^{th} group and fixed observed covariate vector x_{ij}, is assumed as

$$h_{ij}\left(t \mid y_i, x_{ij}\right) = y_i h_0\left(t\right)\exp\left(x_{ij}^{'}\beta\right), i = 1,\ldots,n,\ j = 1,\ldots,n_i \qquad [1]$$

Where $h_0(t)$ an unknown baseline hazard function is common to every subject and β is the vector of fixed effect parameters. The shared frailty variable Y_i is assumed to be independent and identically distributed for groups of patients.

In most commonly to model the frailty, Gamma distribution is used as

$$f\left(y_i\right) = \frac{1}{\Gamma\theta}\theta^{\theta} y_i^{\theta-1}\exp\left(-\theta y_i\right), i = 1,\ldots,n \qquad [2]$$

The higher values of θ^{-1} signify larger variances for y_i, consequently greater heterogeneity among different groups of patients. The role of shared frailty model is more useful when we consider multivariate survival times. The frailty ν is to follow a gamma distribution g (ν; θ). The joint survival function for the k_i individuals within the i^{th} group is

$$S\left(t_{i1},\ldots,t_{ik_i}\right) = \Pr\left(T_{i1} > t_{i1},\ldots,T_{ik_i} > t_{ik_i}\right) \qquad [3]$$

$$= \int_0^{\infty}\prod_{j=1}^{k_i}\Pr\left(T_{ij} > t_{\ ij} \mid \nu_i\right)g\left(\nu_i\right)d\nu_i \qquad [4]$$

$$= \left[1 + \frac{1}{\theta}\sum_{j=1}^{k_i}\Lambda_0\left(t_{ij}\right)\exp\left(\beta^t X_{ij}\right)\right]^{-\theta} \qquad [5]$$

where β, θ, $\Lambda_0(t)$ are estimates.

ISSN: 2278-4632
Vol-12 Issue-01 No.01: 2022

APPLICATION TO BREAST CANCER DATA

The application of accelerated failure time (AFT) with and without frailty models comparisons and shared gamma frailty models are also presented. The Breast cancer data obtained from Government Hospital includes 522 women, diagnosed and treated with adjuvant and neo-adjuvant therapy between January 2000 - December 2008 and follow up period up to May 2010 is used for comparison of different set of covariates with respect to Cancer stage of Share gamma frailty model. The prognostic factors considered for the analysis are, number of children, abortions, menopause status, family history of cancer, cancer stages, radiotherapy treatment, chemotherapy treatment, The event is coded as 1 and censoring is coded as 0,the dependent variable is treatment response under treatment (Raman and Venkatesan, 2012).

The basic step in any parametric regression (AFT) models is to identify whether the dependent variable follows any particular pattern of probabilistic distribution. Hence, an attempt has been made to assess whether the survival time of breast cancer patients follows any known probabilistic distribution. The accelerated failure time (AFT) model is an alternative to the PH model for the analysis of survival data. Under AFT models, we measure the direct effect of the explanatory variables on the survival time instead of hazard, as we do in both PH model. The common probability distributions considered here are exponential, Weibull, log-logistic and log-normal models. The parameter estimates are obtained by assuming the exponential, weibull, log-normal and log-logistic distributions.

Figure 1.1 to 1.4 shows the survival function graph of the five different stages of breast cancer using the exponential, weibull distribution, log-normal and log-logistic distribution treatment response times are nearly similar.

ISSN: 2278-4632
Vol-12 Issue-01 No.01: 2022

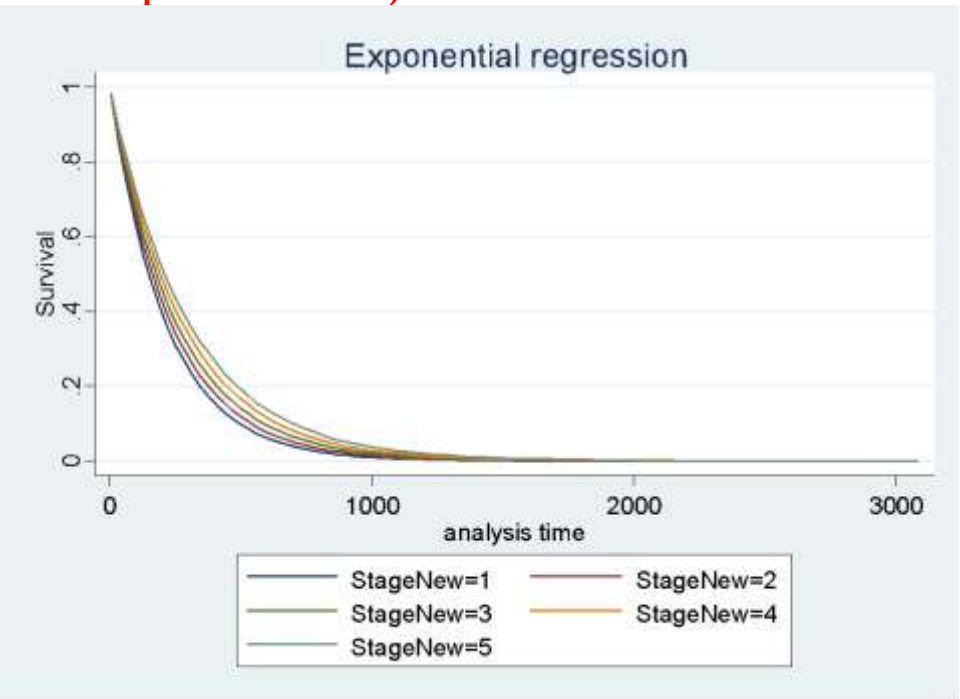

Figure 1.1 Survival function graph for stages with Exponential distribution

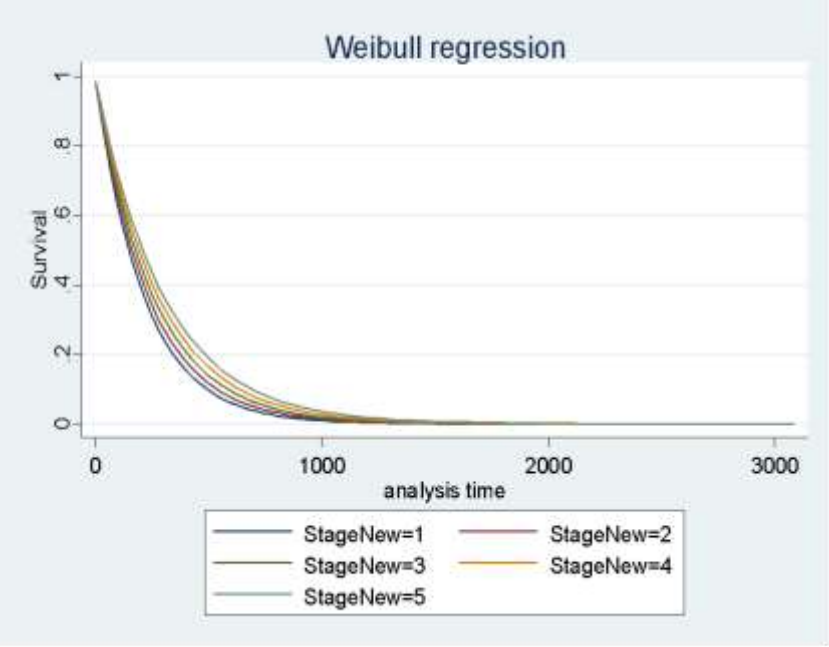

Figure 1.2 Survival function graph for stages with Weibull distribution

ISSN: 2278-4632
Vol-12 Issue-01 No.01: 2022

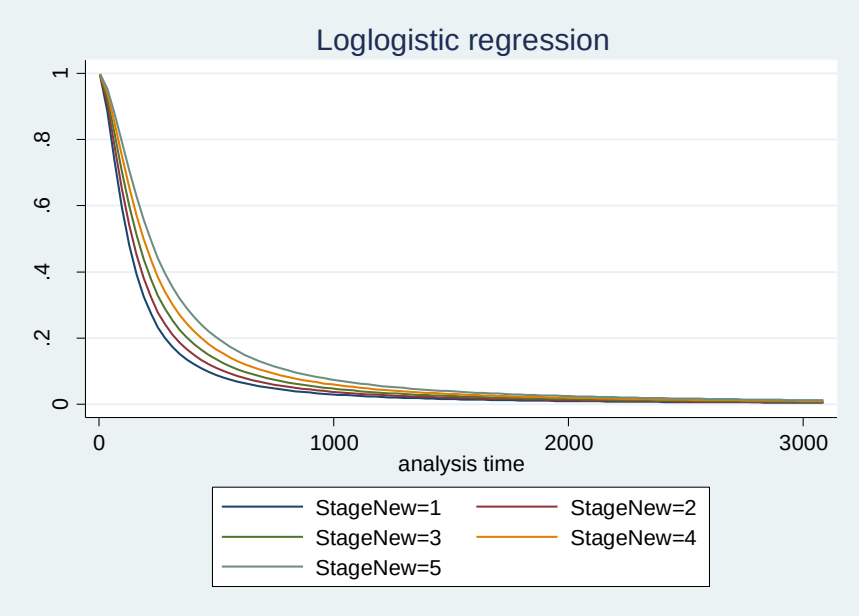

Figure 1.3 Survival function graph for stages with Log-logistic distribution

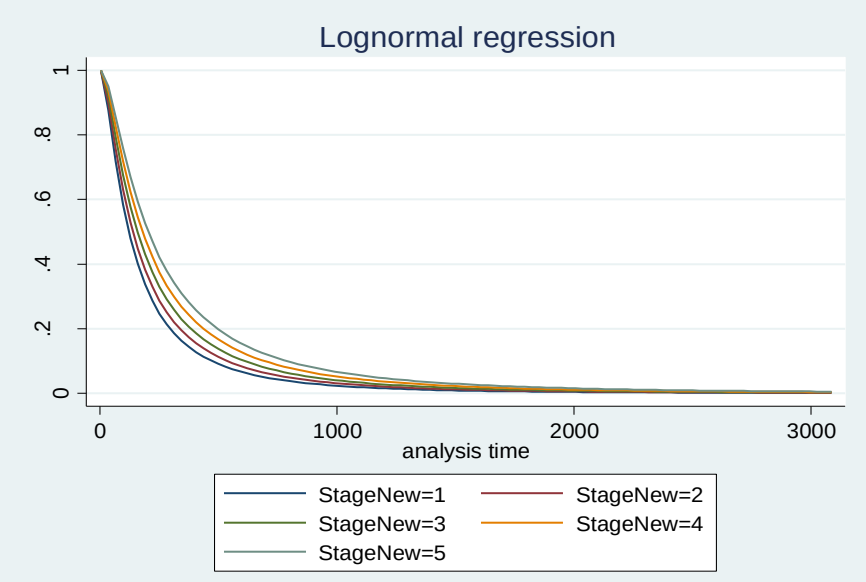

Figure 1.4 Survival function graph for stages with Log-normal distribution

ISSN: 2278-4632
Vol-12 Issue-01 No.01: 2022

Table 1.1 Accelerated Failure Time (AFT) Models

Covariates	Weibull		Exponential		Log-normal		Log-logistic	
	HR	S.E	HR	S.E	Coefficients	S.E	Coefficients	S.E.
Stage	0.8940*	0.0327	0.8970*	0.0327	0.1269*	0.0375	0.1407*	0.0376
Abortions	1.2002	0.1368	1.1802	0.1341	-0.0685	0.1173	-0.0692	0.1102
No. of children	0.9687	0.0240	0.9715	0.0242	0.0143	0.0270	0.0126	0.0271
Chemotherapy	3.0910	0.7916	2.7916	0.6969	-0.4557*	0.2507	-0.3345	0.2720
Radiotherapy	1.1900	0.1152	1.1737	0.1129	-0.07062	0.0954	-0.0683	0.0939
F/H of Cancer	1.0687	0.2401	1.0540	0.2366	0.1009	0.2288	0.0779	0.2163
Menopause	0.9734	0.0953	0.9728	0.0951	0.0750	0.0997	0.0241	0.0990
Deviance	1516.0		1519.5		1488.23		1492.7	

* Significant at 5% level

ISSN: 2278-4632
Vol-12 Issue-01 No.01: 2022

Table 1.2 Accelerated Failure Time with Gamma Frailty Models

Covariates	Weibull		Exponential		Log-normal		Log-logistic	
	HR	S.E	HR	S.E	Coefficients	S.E	Coefficients	S.E.
Stage	0.8299*	0.0447	0.8901*	0.0347	0.1269*	0.0375	0.1407*	0.0376
Abortions	1.1578	0.1767	1.1589	0.1386	-0.0685	0.1173	-0.0692	0.1102
No. of children	0.9714	0.0361	0.9736	0.0263	0.0413	0.0270	0.0126	0.0271
Chemotherapy	2.0867*	0.8167	2.5154*	0.7093	-0.4557*	0.2507	-0.3344	0.2720
Radiotherapy	1.1650	0.1520	1.1633	0.1173	-0.0706	0.0954	-0.0683	0.0939
F/H of Cancer	0.9420	0.2857	1.0228	0.2411	0.1009	0.2288	0.07798	0.2163
Menopause	0.9857	0.1337	0.9747	0.1005	0.0750	0.0997	0.0241	00990
Deviance	1483.97		1516.5		1488.23		1492.65	

* Significant at 5% level

Table 1.1 shows the comparison of different AFT models results. The prognostic factors, stage of the disease were found statistically significant covariates in all the AFT models, except log-normal. In addition to that chemotherapy is significant factor in log-normal. The deviance of log-normal is 1488.23 which is lowest deviance compared with all other AFT models. It's indicates that the log-normal AFT is better model than the others.

The application of accelerated failure time with gamma frailty models to account heterogeneity effects due to unobserved covariates. The AFT model with frailty is identify the presence of any heterogeneity in the model. Hence, an attempt has been made to assess whether the survival time of breast cancer patients follows any known probabilistic distribution, with exponential, Weibull, log-logistic and log- normal with gamma frailty models.

Table 1.2 shows the comparison of different AFT models with gamma frailty results. The prognostic factors stages of the disease and chemotherapy treatment were found statistically significant covariates in all the AFT with gamma frailty models, except log-logistic model. The deviance of Weibull is 1483.97 which are less compared with other gamma frailty models. The Accelerated failure time with gamma frailty model accounts more heterogeneity in Weibull AFT compared with other AFT gamma frailty models. It's indicates that the Weibull AFT with gamma frailty is better fit model for breast cancer data.

The comparison of AFT without and with gamma frailty model results shows the deviance of Weibull is 1483.97 which are less compared with other gamma frailty models, as well as in Weibull model without frailty. The Weibull AFT with gamma frailty model accounts more heterogeneity in compared with other AFT without frailty models and AFT with gamma frailty models. Hence we conclude that Weibull AFT with gamma frailty is better fit model for breast cancer data.

ISSN: 2278-4632
Vol-12 Issue-01 No.01: 2022

Table 1.3 Shared Gamma Frailty Model with respect to Stage

Model No.	Covariate	HR	S.E	95%C.I
Model I	Chemotherapy	1.873*	0.493	(1.120 , 3.142)
	Menopause status	0.954	0.089	(0.793 , 1.147)
	F/H cancer	0.935	0.385	(0.417 , 2.097)
	Theta	0.0508	0.054	
	Deviance	5184 , P = 0.017		
Model II	Chemotherapy	1.943*	0.514	(1.156 , 3.260)
	Menopause status	0.986	0.962	(0.814 , 1.194)
	F/H cancer	0.946	0.390	(0.421 , 2.122)
	No. of Children	0.973	0.024	(0.925 , 1.022)
	Abortions	1.1200	0.135	(0.961 , 1.498)
	Theta	0.0512	0.054	
	-2LL(Deviance)	5050 , P = 0.018		
Model III	Chemotherapy	2.135*	0.575	(1.259 , 3.621)
	Radiotherapy	1.199*	0.114	(0.994 , 1.144)
	Menopause status	0.910	0.109	(0.718 , 1.152)
	F/H cancer	0.965	0.398	(0.429 , 2.169)
	No. of Children	0.967	0.024	(0.920 , 1.017)
	Abortions	1.196	0.136	(0.957 , 1.495)
	Age	1.005	0.005	(0.994 , 1.016)
	Theta	0.038	0.047	
	Deviance	5145.6 , P = 0.049		

Table 1.3 shows the shared gamma frailty model with respect to stages of the diseases. The model I includes the covariates chemotherapy, menopause status, family history of cancer, this model shows chemotherapy is significant factor and deviance is 5184. In model II includes the additional covariates number of children, abortions, and the deviance becomes less 5050, but it reflects the same significant covariates. In model III includes one more additional covariates of age which doesn't contribute to the model, because the deviance value is increased to 5145.6. So model II covariates yields better results for the breast cancer data.

SUMMARY

The proportionality assumption in case of long term follow up may not be always satisfied (Valsecehi et al., 1996). The parametric regression models are applied to overcome this situation. When the survival time has a specified statistical distribution, the statistical power of parametric survival model is higher than non-parametric and semi-parametric models. The effect of the baseline distribution had a great role on the hazards. The AFT model is an alternative method for the analysis of survival data. It is modeled on the survival function directly, with covariates assumed to act multiplicatively directly on the time scale, thus accelerating or decelerating time leading to failure of AFT models. Even when hazards are not proportional based on the asymptotic results, the AFT models would lead to more efficient parameter estimates than Cox model under certain circumferences. The AFT model provides an estimate of the median survival ratios.

The comparison of AFT models shows the significant for the covariates stages in Weibull, Exponential, Log-normal, except Log-normal where chemotherapy is also significant factor. Among the all, the AFT model lowest deviance is show lognormal model which is 1488.23. The log-normal AFT is better model for breast cancer data.

The frailty term varies from individual to individual and is not observable. Hence the distribution of frailty of the population distribution must be specified. Since the hazard function is non-negative, frailty term must be restricted to non-negative values. The role of the choice of gamma frailty distribution effects, it considers ph models with weibul hazards, individual random effects (frailties), whenever detected, it can be made to disappear by elementary model transformation. Random effects model on the other hand, in which group of individuals share some common effect, can be used as being ‘multilevel’ with variation both between and within groups. The logic of comparing the distributional effect is justified for the data by the above arguments. The multivariate frailty model (group factor) assumes unexplained heterogeneity is shared by related individuals and frailty is common to several individuals. In this section, we focus only on univariate case since most of the clinical trial data is on a per subject basis. In most of the cases, a frailty model can only imply a positive correlation within group.

The Accelerated failure time model with gamma frailty and it shows that the covariate stage is significant in exponential and lognormal, whereas chemotherapy is significant in all the models. The deviance of Weibull is 1483.97 which are less compared with other gamma frailty models, as well as in Weibull model without frailty. The Accelerated failure time with gamma frailty model accounts more heterogeneity in Weibull distribution compared with other

Accelerated failure time model. The Weibull AFT with gamma frailty is better model and also heterogeneity is accounted in stages of disease and chemotherapy treatment. The shared gamma frailty model with respect to stage, model II covariates accounts more heterogeneity compared with other models I and III.

REFERENCES:

[1] Aalen, O.O., (1988): Heterogeneity in Survival Analysis. Statistics in Medicine. 7,121-37.

[2] Cox, D.R., Oakes, D., (1984): Analysis of Survival Data. London Chapman and Hall.

[3] Duncan, Kerr G.R., (1976): The curability of breast cancer. Br Med J. 2, 781–783.

[4] Haybittle, J.L., (1959): The estimation of the proportion of patients cured after treatment for Cancer of the breast. Br J Radiol. 32, 725–733.

[5] Hosmer, D.W., Lemeshow S., (1999): Applied Survival Analysis: regression modeling of time to event data. New York, Willey.

[6] Keiding, N., Anderson, P.K., and Klein J.P., (1997): The role of frailty models and accelerated failure time models in describing heterogeneity due to omitted covariates. Statistics in Medicine.16, 215-24.

[7] Klein Baum, D.G., (1996): Survival analysis: a self learning text, New York, Springer.

[8] Lee E.T., and Wang J.W., (2003): Statistical Methods for survival Data Analysis. New Jersey, USA: John Willey.

[9] Maggard M.A., Thompson J.E., and KO C.Y., (2003). Why do breast cancer mortality rates vary across states? Am Surg .69, 59–62.

[10] Ponnuraja, C., and Venkatesan, P., (2010): Survival Models for exploring tuberculosis clinical trial data-an empirical comparison. Indian J Sci. Techno.3, 755-758.

[11] Rajaeefard, A.R., Baneshi, M.R., Talei, A.R., and Mehrabani, D., (2009): Survival Models in Breast Cancer Patients. Iranian Red Crescent Medical Journal. 11,295-300.

[12] Rama, R., Swaminathan, R., and Venkatesan, P., (2010): Cure models for estimating hospital-based breast cancer survival, Asian Pac J Cancer Prev. 11,387-91.

[13] Wei, L.J., (1992): The Accelerated failure time model: A useful alternative to the Cox regression model in survival analysis (with discussion). Statistics in Medicine.11, 1871-79.

[14] Weipan., (2001): Using frailties in the accelerated failure time model. Life time data analysis, 7, 55-64.

[15] Zahl, P.H., Tretli, S., (1997): Long-term survival of breast cancer in Norway by age and clinical stage. Stat Med, 16, 1435–1449.

