Journal of Statistical Research *ISSN 0256 - 422 X*
2006, Vol. 40, No. 1, pp. 45-57
Bangladesh

BAYESIAN INFERENCE USING BURR MODEL UNDER ASYMMETRIC LOSS FUNCTION: AN APPLICATION TO CARCINOMA SURVIVAL DATA

Abdus S. Wahed
Department of Biostatistics, University of Pittsburgh Graduate School of Public Health, Pittsburgh, Pennsylvania 15261, U.S.A.
Email: wahed@pitt.edu

SUMMARY

We present Bayes estimators for the parameters of Burr type XII distribution under the symmetric squared error loss function and the asymmetric linear-exponential loss function based on a simple prior distribution. In all cases the estimator turns out to be ratios of integrals. We present approximate Bayes estimators based on different approximation techniques. We demonstrate the application of Burr type XII distribution to model the time to death in a clinical trial of carcinoma patients comparing two therapies.

Keywords and phrases: Bayesian Inference, LINEX Loss Function, Survival Analysis.

AMS Classification: 62F15, 62N01, 62N02, 62N03, 62N05

1 Introduction

Introduced by Burr (1942), Burr distribution of Type XII with the distribution function

$$F(x;\alpha,\beta) = 1 - (1 + x^{\beta})^{-\alpha}, x > 0, \alpha, \beta > 0, \tag{1.1}$$

yields a wide range of values of skewness and kurtosis and has been used as probability models in various applications. Rodriguez (1977) gives a comprehensive overview of Burr type XII distribution. From the perspective of inferential aspects of the parameters α and β, frequentist approach has drawn more attention than the Bayesian approach. As an estimator, a natural choice that the users mostly made is the maximum likelihood estimator (MLE), mostly because of its ease of use. Wingo (1983, 1993) developed mathematical and computational methodology for ML fitting of the Burr XII distribution to censored life test data. A Bayesian approach to inference about the parameters of a Burr distribution was taken by Papadopoulos (1978) where he used the distribution as a failure model. Some theoretical results on Bayesian inference on Burr parameters, reliability and hazard rate are given in AL-Hussaini and Jaheen (1994). Soliman (2002) considered Bayes estimators

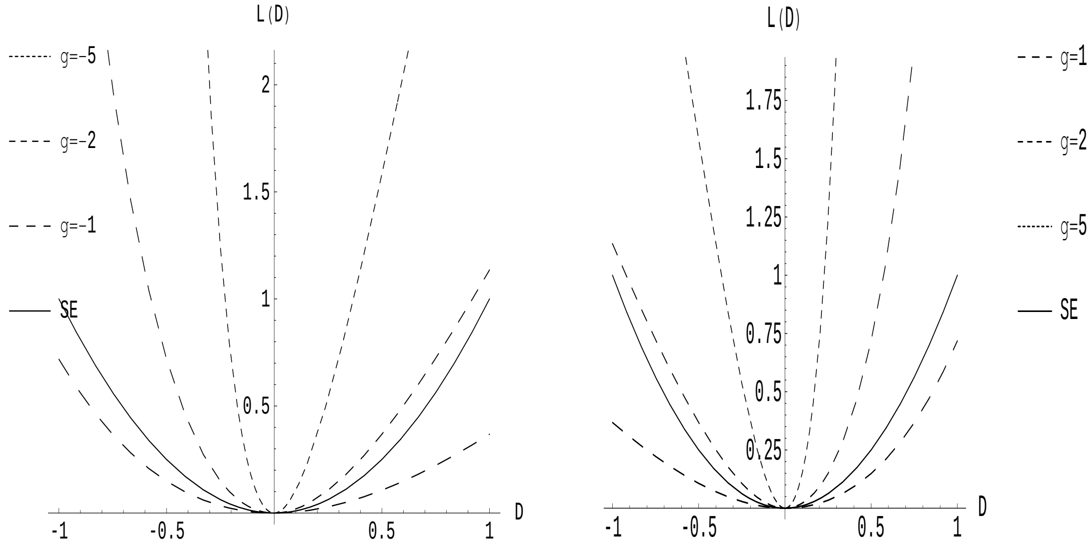

Figure 1: LINEX and squared error (SE) loss functions plotted as a function of the error $D = \hat{\theta} - \theta$: left panel for $\gamma < 0$ and right panel for $\gamma > 0$.

of reliability function in a generalized life-model that included the Burr XII distribution. Soliman (2005) extended their work to incorporate progressively right censored data.

Use of symmetric loss functions, such as squared error loss is common in Bayesian analysis. In most cases it is done for convenience but may not be appropriate in many real life situations (Varian, 1975). For example, in survival analysis, an overestimation would create false optimism and an underestimation might lead to depression and hence deteriorate the health condition. Thus, it is of importance to consider Bayes estimators under asymmetric loss functions such as linear-exponential (LINEX) family of loss functions introduced first by Varian (1975) and later generalized by Zellner (1986). The LINEX loss function for estimating a p-dimentional vector θ by $\hat{\theta}$ is defined as follows:

$$L(\hat{\theta}, \theta) = \sum_{k=1}^{p} \omega_k \left[e^{\gamma_k(\hat{\theta}_k - \theta_k)} - \gamma_k(\hat{\theta}_k - \theta_k) - 1 \right], \; \omega_k > 0, \; \gamma_k \neq 0; \; k = 1, 2, \cdots p. \qquad (1.2)$$

where $\hat{\theta}_k$ and θ_k are the k^{th} components of $\hat{\theta}$ and θ respectively. Figure 1 depicts the nature of the LINEX loss function in comparison with the squared error loss function.

Apart from Soliman (2002, 2005), all other studies used only quadratic loss functions. However, these studies failed to provide a case study where the application of proposed methods would be appropriate. Additionally, these two articles employed only Lindley's approximation (Lindley, 1980) to obtain approximate Bayes estimators. Tierney and Kadane (1986) has shown that their method provides better approximation to the Bayes estimators than Lindley's as the error in approximation for the latter is of order $O(n^{-1})$ as compared with the $O(n^{-2})$ of the former. In this paper we have considered the joint Bayes estimation of (α, β) under both squared error (SE) and LINEX loss functions and used both approximation techniques to approximate them. Although AL-Hussaini and Jaheen (1994) have used both approximation techniques, the prior they have used is more complicated and needs expert judgment in guessing the hyperparameters.

The purpose of this study is two-fold. Facilitation of functions needed to approximate Bayes estimators for Burr parameters using a simple prior is our first goal, followed by the demonstration of an application to a real dataset. The article is organized as follows. In Section 2, we specify the prior and compute the joint posterior distribution for the parameter vector (α, β) . The expressions for the Bayes estimators under LINEX and SE loss functions and the corresponding approximations are given in Section 3. An application to a survival dataset has been considered in Section 4. We wrap up with a discussion in Section 5.

2 Prior, Likelihood and Posterior

In this study, we have assumed, for simplicity, that the two parameters α and β are stochastically independent. That is, their joint prior distribution can be constructed as the product of the marginal priors. We assume that the marginal prior distribution of α is given by

$$\pi(\alpha) \propto \frac{1}{\alpha},\ \alpha > 0. \tag{2.1}$$

whereas the marginal prior distribution for β is assumed to follow a $Gamma(\psi, \phi)$ distribution. So that the joint prior for (α, β) can be written as

$$\pi(\alpha, \beta) \propto \frac{\phi^{\psi}}{\alpha\Gamma(\psi)} e^{-\phi\beta}\beta^{\psi-1},\ \alpha, \beta \geq 0; \psi, \phi > 0. \tag{2.2}$$

The prior specified here is different from the more general one specified in AL-Hussaini and Jaheen (1994) in the sense that the latter allowed the dependence of one parameter on the other through a conditional distribution of α given β. Our prior employs only two hyperparameters compared to the four specified for the prior in AL-Hussaini and Jaheen (1994).

Let $\mathbf{x} = (x_1, x_2, ..., x_n)$ be n independent observations from the density (1.1). Then the likelihood function for (α, β) given the data $\mathbf{x}$ is obtained as

$$\ell(\alpha, \beta|\mathbf{x}) = exp\left[nln\alpha + nln\beta + (\beta - 1)S_1(\mathbf{x}) - (1+\alpha)S_2(\beta; \mathbf{x})\right], \tag{2.3}$$

where $S_1(\mathbf{x}) = \sum_{i=1}^{n} lnx_i$ and $S_2(\beta; \mathbf{x}) = \sum_{i=1}^{n} ln(1 + x_i^{\beta})$.

A combination of the prior (2.2) and the likelihood (2.3) produces the posterior density of (α, β) of the form

$$P(\alpha, \beta|\mathbf{x}) = \{g(\mathbf{x})\}^{-1} exp\left[(n-1)ln\alpha + (n+\psi-1)ln\beta + \{S_1(\mathbf{x}) - \phi\}\beta - (1+\alpha)S_2(\beta; \mathbf{x})\right], \tag{2.4}$$

where the function $g(\mathbf{x})$ is such that the expression in (2.4) is a proper density of α and β. Integrating (2.4) with respect to (α, β) through iterative integrals, we obtain

$$g(\mathbf{x}) = \Gamma(n) \int_0^\infty exp\left[(n+\psi-1)ln\beta + \{S_1(\mathbf{x}) - \phi\}\beta - S_2(\beta;\mathbf{x})\right]/S_2^n(\beta;\mathbf{x})d\beta.$$

The posterior distribution (2.4) involves an integral in the denominator which cannot be evaluated in closed form. Accordingly, the Bayes estimators of α and β cannot be expressed in any simple form and application of approximation techniques is essential in order to evaluate them for a given data set. For the chosen priors, it has been observed that the posterior is dominated by the likelihood, not the prior. Hence, the inferences made from this posterior will rely more on data than prior belief.

3 Bayes estimators

Bayes estimator of a parameter under squared error loss function is just the posterior mean. From the posterior distribution of the parameters (α, β) (2.4), the Bayes estimators of α under squared error loss is given by

$$\hat{\alpha}_{SE} = E(\alpha|\mathbf{x}) = \frac{n\int_0^\infty exp\left[(n+\psi-1)ln\beta + (S_1(\mathbf{x}) - \phi)\beta - S_2(\beta;\mathbf{x})\right]/S_2^{n+1}(\beta;\mathbf{x})d\beta}{\int_0^\infty exp\left[(n+\psi-1)ln\beta + (S_1(\mathbf{x}) - \phi)\beta - S_2(\beta;\mathbf{x})\right]/S_2^{n}(\beta;\mathbf{x})d\beta}, \tag{3.1}$$

and that of β by

$$\hat{\beta}_{SE} = E(\beta|\mathbf{x}) = \frac{\int_0^\infty exp\left[(n+\psi)ln\beta + (S_1(\mathbf{x}) - \phi)\beta - S_2(\beta;\mathbf{x})\right]/S_2^{n}(\beta;\mathbf{x})d\beta}{\int_0^\infty exp\left[(n+\psi-1)ln\beta + (S_1(\mathbf{x}) - \phi)\beta - S_2(\beta;\mathbf{x})\right]/S_2^{n}(\beta;\mathbf{x})d\beta}. \tag{3.2}$$

Under the loss function (1.2) the Bayes estimator $\hat{\theta}_k$ of θ_k is given by

$$\hat{\theta}_k = -\frac{1}{\gamma_k} lnE\left(e^{-\gamma_k\theta_k}|\mathbf{x}\right) \tag{3.3}$$

where $E(.|\mathbf{x})$ stands for the posterior expectation given the sample data $\mathbf{x}$. Using this, the Bayes estimator of α from the joint posterior (2.4) under LINEX loss is given by

$$\hat{\alpha}_{LINEX} = -\frac{1}{\gamma_1}\ln\left[\frac{\int_0^\infty exp\left[(n+\psi-1)ln\beta + (S_1(\mathbf{x}) - \phi)\beta - S_2(\beta;\mathbf{x})\right]/\{\gamma_1 + S_2(\beta;\mathbf{x})\}^n d\beta}{\int_0^\infty exp\left[(n+\psi-1)ln\beta + (S_1(\mathbf{x}) - \phi)\beta - S_2(\beta;\mathbf{x})\right]/S_2^{n}(\beta;\mathbf{x})d\beta}\right]. \tag{3.4}$$

Similarly, the Bayes estimator of β under LINEX loss function (1.2) is obtained as,

$$\hat{\beta}_{LINEX} = -\frac{1}{\gamma_2}\ln\left[\frac{\int_0^\infty exp\left[(n+\psi-1)ln\beta + (S_1(\mathbf{x}) - \phi - \gamma_2)\beta - S_2(\beta;\mathbf{x})\right]/S_2^{n}(\beta;\mathbf{x})d\beta}{\int_0^\infty exp\left[(n+\psi-1)ln\beta + (S_1(\mathbf{x}) - \phi)\beta - S_2(\beta;\mathbf{x})\right]/S_2^{n}(\beta;\mathbf{x})d\beta}\right] \tag{3.5}$$

All four estimators (3.1),(3.2),(3.4)and (3.5) are in the form of ratios of two integrals - none of which can be simplified further to closed forms. We apply two different approximation methods to evaluate the Bayes estimators of α and β. The first approximation technique due to Lindley (1980) uses Taylor's series expansion of the integral expression around maximum likelihood estimator. The second is due to Tierney and Kadane(1986). These methods are known to work well for the cases where the dimensionality of the parameter vector is small. For a detailed and useful description of these techniques we refer the readers to Press (2003). The other approximation techniques are based mainly on Monte-Carlo integration which are time consuming and are more applicable to higher dimensional problems.

To approximate the Bayes estimators using Lindley's approximation (Appendix A, Equation (A.3)), we introduce the following notations:

$$S_3(\beta;\mathbf{x}) = \sum_{i=1}^n \ln x_i(1+x_i^\beta)^{-1}x_i^\beta = \delta S_2(\beta;\mathbf{x})/\delta\beta,$$
$$S_4(\beta;\mathbf{x}) = \sum_{i=1}^n (\ln x_i)^2(1+x_i^\beta)^{-2}x_i^\beta = \delta S_3(\beta;\mathbf{x})/\delta\beta, \text{ and}$$
$$S_5(\beta;\mathbf{x}) = \sum_{i=1}^n (\ln x_i)^3(1+x_i^\beta)^{-3}x_i^\beta(1-x_i^\beta) = \delta S_4(\beta;\mathbf{x})/\delta\beta.$$

With these new definitions, in terms of the notations used in Equation (A.3), the first order derivatives are $\rho_\alpha = \alpha^{-1}$, $\rho_\beta = -\psi + \beta^{-1}(\phi - 1)$, $L_\alpha = n/\alpha - S_2(\beta;\mathbf{x})$, $L_\beta = n/\beta + S_1(\mathbf{x}) - (1+\alpha)S_3(\beta;\mathbf{x})$. Similarly, the second order derivatives are $L_{\alpha\alpha} = -\frac{n}{\alpha^2}$, $L_{\alpha\beta} = -S_3(\beta;\mathbf{x})$, $L_{\beta\beta} = -\frac{n}{\beta^2} - (1+\alpha)S_4(\beta;\mathbf{x})$. The third order derivatives are then calculated as follows: $L_{\alpha\alpha\alpha} = \frac{2n}{\alpha^3}$, $L_{\alpha\beta\alpha} = L_{\alpha\alpha\beta} = L_{\beta\alpha\alpha} = 0$, $L_{\alpha\beta\beta} = L_{\beta\alpha\beta} = L_{\beta\beta\alpha} = -S_4(\beta;\mathbf{x})$, $L_{\beta\beta\beta} = \frac{2n}{\beta^3} - (1+\alpha)S_5(\beta;\mathbf{x})$. The σ-components are then $\sigma_{\alpha\alpha} = -\frac{1}{L_{\alpha\alpha}}$, $\sigma_{\beta\beta} = -\frac{1}{L_{\beta\beta}}$.

The maximum likelihood estimators for α and β, $\hat\alpha$ and $\hat\beta$ satisfy $L_\alpha = 0 = L_\beta$. Explicitly, $\hat\alpha = nS_2^{-1}(\hat\beta;\mathbf{x})$ and $\hat\beta$ is obtained as a solution to the equation

$$n\hat\beta^{-1} + S_1(\mathbf{x}) - (1 + nS_2^{-1}(\hat\beta;\mathbf{x}))S_3(\hat\beta;\mathbf{x}) = 0.$$

The parametric functions, for instance, $L_{\beta\beta\beta}$, when parameters are substituted for the corresponding MLE's are denoted by putting a $\hat{()}$ on it. For instance,

$$\hat L_{\beta\beta\beta} = L_{\beta\beta\beta}|_{\alpha=\hat\alpha,\beta=\hat\beta} = \frac{2n}{\hat\beta^3} - (1+\hat\alpha)S_5(\hat\beta;\mathbf{x}).$$

Applying Lindley's approximation technique and using the notations defined above, we obtain approximate Bayes estimator of α

$$\hat\alpha_{SE}^{Lindley} = \hat\alpha - \frac{\hat\alpha^2\hat\beta^2 S_4(\hat\beta;\mathbf{x})/2n}{n + (1+\hat\alpha)\hat\beta^2 S_3(\hat\beta;\mathbf{x})}. \tag{3.6}$$

Similar derivation gives the following estimator for β under Lindley's approximation:

$$\hat\beta_{SE}^{Lindley} = \hat\beta + \frac{\hat\beta^{-1}(\psi - 1) - \phi}{n\hat\beta^{-2} + (1+\hat\alpha)S_4(\hat\beta;\mathbf{x})} + \frac{n\hat\beta^{-3} - (1+\hat\alpha)S_5(\hat\beta;\mathbf{x})/2}{\{n\hat\beta^{-2} + (1+\hat\alpha)S_4(\hat\beta;\mathbf{x})\}^2}. \tag{3.7}$$

Following the same procedure, an application of Lindley's approximation theorem to the integral expression of (3.4) gives

$$\hat{\alpha}_{LINEX}^{Lindley} = \hat{\alpha} - \frac{1}{\gamma_1} \ln \left[1 + \frac{\gamma_1 \hat{\alpha}^2}{2n} \right]. \tag{3.8}$$

Similarly,

$$\hat{\beta}_{LINEX}^{Lindley} = \hat{\beta} - \frac{1}{\gamma_2} \ln \left[1 + \frac{\gamma_2}{2} \left\{ \frac{1 + 2(\phi - \hat{\beta}^{-1}(\psi - 1))}{n\hat{\beta}^{-2} + (1+\hat{\alpha})S_4(\hat{\beta}; \mathbf{x})} - \frac{2n\hat{\beta}^{-3} - (1+\hat{\alpha})S_5(\hat{\beta}; \mathbf{x})}{\{n\hat{\beta}^{-2} + (1+\hat{\alpha})S_4(\hat{\beta}; \mathbf{x})\}^2} \right\} \right]. \tag{3.9}$$

It is not difficult to investigate the fact that these estimators are asymptotically equivalent to the corresponding maximum likelihood estimators.

To explain the Tierny and Kadane's method, let us first consider estimating α under LINEX loss function (1.2). A brief description of this approximation method is given in Appendix B. The posterior expectation that is of interest is $E\left(e^{-\gamma_1 \alpha}|\mathbf{x}\right)$. Let, $u(\alpha, \beta) = e^{-\gamma_1 \alpha}$, $l(\alpha, \beta) = n^{-1}\{L(\alpha, \beta; x) + \rho(\alpha, \beta)\}$, and $l^*(\alpha, \beta) = n^{-1} \ln u(\alpha, \beta) + l(\alpha, \beta)$. Also let $(\hat{\alpha}_l, \hat{\beta}_l)$ and $(\hat{\alpha}_{l^*}, \hat{\beta}_{l^*})$ be the values of (α, β) at which the functions l and l^* respectively attains their maximum. That is, $(\hat{\alpha}_l, \hat{\beta}_l) = \arg\max_{(\alpha,\beta)} l(\alpha, \beta)$, and $(\hat{\alpha}_{l^*}, \hat{\beta}_{l^*}) = \arg\max_{(\alpha,\beta)} l^*(\alpha, \beta)$. In our case,$(\hat{\alpha}_l, \hat{\beta}_l)$ is given by the solutions of the equations

$$\begin{aligned} \hat{\beta}_l^{-1}(n + \psi - 1) + S_1(\mathbf{x}) - \phi - (1 + \hat{\alpha}_l)S_3(\hat{\beta}_l; \mathbf{x}) &= 0 \\ n - 1 - \hat{\alpha}_l S_2(\hat{\beta}_l; \mathbf{x}) &= 0 \end{aligned}$$

and $(\hat{\alpha}_{l^*}, \hat{\beta}_{l^*})$ by

$$\begin{aligned} \hat{\beta}_{l^*}^{-1}(n + \psi - 1) + S_1(\mathbf{x}) - \phi - (1 + \hat{\alpha}_{l^*})S_3(\hat{\beta}_{l^*}; \mathbf{x}) &= 0 \\ n - 1 - \gamma_1 \alpha_{l^*} - \hat{\alpha}_{l^*} S_2(\hat{\beta}_{l^*}; \mathbf{x}) &= 0 \end{aligned}$$

Further let $\Sigma(\alpha, \beta)$ and $\Sigma^*(\alpha, \beta)$ are negatives of inverse Hessians of $l(\alpha, \beta)$ and $l^*(\alpha, \beta)$ respectively. In this case,

$$\Sigma^{-1}(\alpha, \beta) = n^{-1} \begin{bmatrix} \frac{n-1}{\alpha^2} & S_3(\beta; \mathbf{x}) \\ S_3(\beta; \mathbf{x}) & \frac{(n+\psi-1)}{\beta^2} + (1+\alpha)S_4(\beta; \mathbf{x}) \end{bmatrix} = \Sigma^{*-1}(\alpha, \beta) \tag{3.10}$$

Hence,

$$|\Sigma(\alpha, \beta)| = |\Sigma^*(\alpha, \beta)| = n^2 \left[\frac{n-1}{\alpha^2} \left\{ \frac{(n+\psi-1)}{\beta^2} + (1+\alpha)S_4(\beta; \mathbf{x}) \right\} - S_3^2(\beta; \mathbf{x}) \right]^{-1}.$$

Then, the Tierny-Kadane approximation to the Bayes estimator of α under LINEX loss function is given by

$$\hat{\alpha}_{LINEX}^{T-K} = -\frac{1}{\gamma_1} \ln I(\mathbf{x}), \tag{3.11}$$

Table 1: Functional expressions for Tierney-Kadane Approximations of Bayes estimators $\hat{\alpha}_{SE}$, $\hat{\beta}_{LINEX}$ and $\hat{\beta}_{SE}$

Function	Expressions for approximating $\hat{\alpha}_{SE}$
$u(\alpha,\beta)$	α
$\lvert\Sigma^*(\alpha,\beta)\rvert$	$n^2\left[\frac{n}{\alpha^2}\left\{\frac{(n+\psi-1)}{\beta^2}+(1+\alpha)S_4(\beta;\mathbf{x})\right\}-S_3^2(\beta;\mathbf{x})\right]^{-1}$
Estimating equations	$n-\hat{\alpha}_{l^*}S_2(\hat{\beta}_{l^*};\mathbf{x})=0$
for $(\hat{\alpha}_{l^*},\hat{\beta}_{l^*})$	$\hat{\beta}_{l^*}^{-1}(n+\psi-1)+S_1(\mathbf{x})-\phi-(1+\hat{\alpha}_{l^*})S_3(\hat{\beta}_{l^*};\mathbf{x})=0$
	Expressions for approximating $\hat{\beta}_{LINEX}$
$u(\alpha,\beta)$	$e^{-\gamma_2\beta}$
$\lvert\Sigma^*(\alpha,\beta)\rvert$	$n^2\left[\frac{n-1}{\alpha^2}\left\{\frac{(n+\psi-1)}{\beta^2}+(1+\alpha)S_4(\beta;\mathbf{x})\right\}-S_3^2(\beta;\mathbf{x})\right]^{-1}$
Estimating equations	$n-1-\hat{\alpha}_{l^*}S_2(\hat{\beta}_{l^*};\mathbf{x})=0$
for $(\hat{\alpha}_{l^*},\hat{\beta}_{l^*})$	$-\gamma_2+\hat{\beta}_{l^*}^{-1}(n+\psi-1)+S_1(\mathbf{x})-\phi-(1+\hat{\alpha}_{l^*})S_3(\hat{\beta}_{l^*};\mathbf{x})=0$
	Expressions for approximating $\hat{\beta}_{SE}$
$u(\alpha,\beta)$	β
$\lvert\Sigma^*(\alpha,\beta)\rvert$	$n^2\left[\frac{n-1}{\alpha^2}\left\{\frac{(n+\psi)}{\beta^2}+(1+\alpha)S_4(\beta;\mathbf{x})\right\}-S_3^2(\beta;\mathbf{x})\right]^{-1}$
Estimating equations	$n-1-\hat{\alpha}_{l^*}S_2(\hat{\beta}_{l^*};\mathbf{x})=0$
for $(\hat{\alpha}_{l^*},\hat{\beta}_{l^*})$	$\hat{\beta}_{l^*}^{-1}(n+\psi)+S_1(\mathbf{x})-\phi-(1+\hat{\alpha}_{l^*})S_3(\hat{\beta}_{l^*};\mathbf{x})=0$

where $I(\mathbf{x})$ is given by the equation (B.1). Similar procedure can be followed to approximate the Bayes estimators of α under SE loss function and Bayes estimators of β under LINEX and SE loss functions. Notice that the function l and hence $(\hat{\alpha}_l,\hat{\beta}_l)$ and Σ does not change no matter what function of the parameters is being estimated. In Table 1 we give the expressions for the component functions of (3.11) and the estimating equations to be solved to obtain corresponding $(\hat{\alpha}_{l^*},\hat{\beta}_{l^*})$. Solving the estimating equations and consequently using the functional forms from Table 1 in Equation (3.11), we approximated the other three estimators (3.2),(3.4)and (3.5) and respectively denoted them by $\hat{\alpha}_{SE}^{T-K}$, $\hat{\beta}_{LINEX}^{T-K}$ and $\hat{\beta}_{SE}^{T-K}$.

4 Data Analysis

We applied our techniques to a subset of carcinoma data set from a clinical trial carried out by Radiation Therapy Oncology Group (Dataset II, Appendix A , Kalbfleisch and Prentice, 2003). Although some patients have been censored, we considered them as complete in our analysis and hence the mean survival estimates presented here will be less than those, had we considered a likelihood based on the censored data. This will be true for all the

Table 2: Bayes estimates of α, β and mean (in years) for the carcinoma data

				Lindley		**Tierney and Kadane**		
Treatment	Parameter	MLE	SE	LINEX	LINEX	SE	LINEX	LINEX
	γ			-2	-5		-2	-5
				$\psi = 2, \phi = 1$				
Radiation Therapy	α	0.78	0.78	0.78	0.79	0.79	0.80	0.81
	β	2.24	2.23	2.22	2.22	2.19	2.24	2.31
	Mean	2.01	2.02	2.01	1.97	2.01	1.92	1.80
Radiation + Chemotherapy	α	1.01	1.01	1.02	1.03	1.02	1.03	1.05
	β	1.96	1.96	1.95	1.95	1.93	1.96	2.02
	Mean	1.58	1.59	1.57	1.54	1.59	1.53	1.46
				$\psi = 5, \phi = 2$				
Radiation Therapy	α	0.78	0.78	0.78	0.79	0.81	0.82	0.83
	β	2.24	2.24	2.23	2.23	2.12	2.15	2.22
	Mean	2.01	2.01	1.99	1.96	2.04	1.95	1.83
Radiation + Chemotherapy	α	1.01	1.01	1.02	1.03	1.03	1.04	1.06
	β	1.96	1.97	1.96	1.95	1.88	1.91	1.96
	Mean	1.58	1.58	1.56	1.53	1.60	1.55	1.47
				$\psi = 5, \phi = 5$				
Radiation Therapy	α	0.78	0.78	0.78	0.79	0.81	0.81	0.83
	β	2.24	2.17	2.16	2.14	2.13	2.17	2.23
	Mean	2.01	2.10	2.09	2.07	2.04	1.95	1.83
Radiation + Chemotherapy	α	1.01	1.01	1.02	1.03	1.03	1.04	1.06
	β	1.96	1.90	1.89	1.88	1.89	1.92	1.97
	Mean	1.58	1.63	1.62	1.59	1.60	1.55	1.47
				$\psi = 10, \phi = 5$				
Radiation Therapy	α	0.78	0.78	0.78	0.79	0.83	0.84	0.85
	β	2.24	2.22	2.21	2.21	2.02	2.04	2.10
	Mean	2.01	2.03	2.02	1.98	2.08	2.01	1.88
Radiation + Chemotherapy	α	1.01	1.01	1.02	1.03	1.05	1.06	1.08
	β	1.96	1.96	1.95	1.95	1.81	1.83	1.88
	Mean	1.58	1.59	1.57	1.54	1.63	1.58	1.50

estimators for both treatment arms. The objective is to compare the mean survival time and survival distributions under two treatments. We demonstrate how, using the estimators derived in previous sections, one can compare the two treatments in question.

In our analysis we have included 195 patients who have been treated for squamous cell carcinoma with either standard radiation therapy or radiation therapy together with a chemotherapeutic agent. The variable of interest here is the survival time, as usual defined by the time between entry into the study and the death. We expressed the time in years for simplicity. We calculated approximate Bayes estimators for the parameters α and β and the corresponding survival mean and survival curves from Burr model for both radiation therapy (standard) and radiation+chemotherapy (test). We considered only negative values of γ since we believe, in estimating survival, underestimation results in more penalty than over estimation. We considered two values for γ, namely, $\gamma = -2$ and $\gamma = -5$.

Table 2 shows the estimates of parameters and mean survival time under different estimation methods. We considered several combinations of hyperparameters to investigate their sensitivity to the estimates. It is observed that the estimates of parameters and mean survival time are barely sensitive to the choice of the prior parameters. For instance, the Bayes estimate of the mean survival time for the radiation therapy group under LINEX loss function with parameter $\gamma = -2$ approximated by Tierney and Kadane method takes values 1.92, 1.95, 1.95, 2.01 for prior parameter combinations $(\psi, \phi) = (2, 1), (5, 2), (5, 5), (10, 5)$ respectively. Thus the proposed prior can be used without paying much attention to the choice of prior hyperparameters. Any reasonable choice such as the ones presented in Table 2 will lead to similar conclusion. In cases, where choice of hyperparameters is a concern, one can use empirical Bayes methods to obtain better guesses for the hyperparameters. Since our choice of prior is such that the prior parameters had little impact on the posterior and for that reason, on the Bayes estimators, we do not pursue it here.

In practice, the choice of γ is dictated by the relative consequence of underestimation and overestimation. In our case of modeling survival data, we picked two such values of γ to reflect the mild and severe consequences of underestimation compared to overestimation. From the results in Table 2, it is also clear that the choice of the LINEX loss function parameter γ had very little impact on the Bayes estimates. In most cases, the Bayes estimates of parameters and the mean survival times were identical to the second decimal places for two parameter values $\gamma = -2$ and $\gamma = -5$.

All Bayes estimators, approximated by Lindley's method or Tierney and Kadane's method, gave rise to similar estimates. Although, in theory, Bayes estimators approximated by Tierney and Kadane's method are supposed to be more accurate than those approximated by Lindley's method, it is interesting to see that for this dataset, for the selected prior and loss function parameters, they provide estimates which are very close to the corresponding maximum likelihood estimate.

The test treatment radiation + chemotherapy seems to have no advantage over the standard radiation therapy as seen by the lower mean survival (approx. 1.6 years) within the test group compared to the standard group (approx. 2.0 years), no matter which method

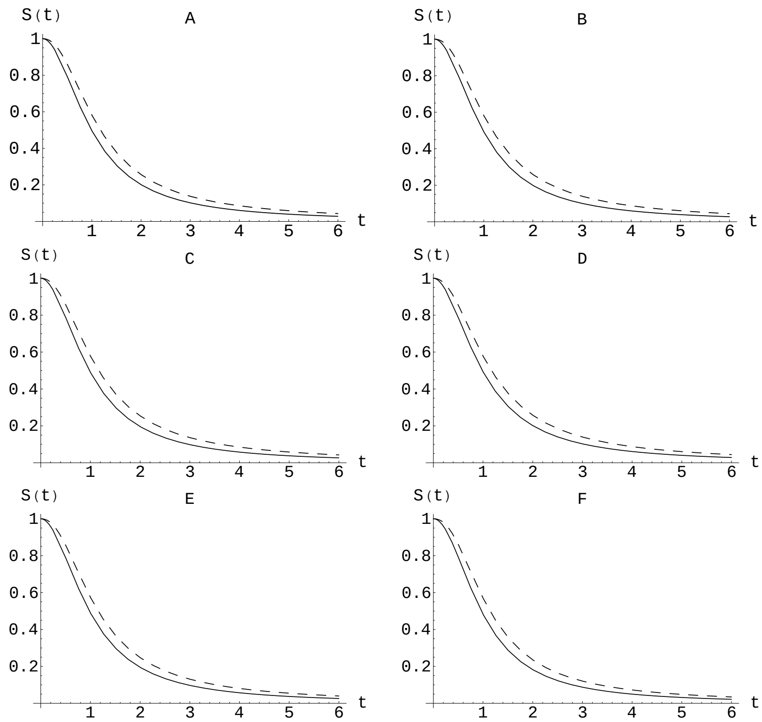

Figure 2: Estimated survival probability curves for carcinoma data for prior parameters $\phi = 1$ and $\psi = 2$ under different estimation scheme (dotted line for radiation therapy and solid line for radiation + chemotherapy): A. Maximum Likelihood (same as SE with Lindley's approximation). B. LINEX under Lindley's with $\gamma = -2$. C. LINEX under Lindley's with $\gamma = -5$. D. SE under Tierney-Kadane. E. LINEX under Tierney-Kadane with $\gamma = -2$. F. LINEX under Tierney-Kadane with $\gamma = -5$.

we use for estimation. We have plotted the estimated survival curves under different methods in Figure 2. Since the estimates of parameters were very similar to each other the graphs were also similar across different methods. The survival curve for the standard radiation therapy group dominates that of radiation + chemotherapy group providing further evidence of ineffectiveness of radiation followed by chemotherapy in extending the survival.

5 Conclusion

We have derived approximate Bayes estimators for the parameters of a Burr type XII distribution using several approximation techniques under symmetric (squared error) and asymmetric (linear-exponential) loss functions. The expression for approximate Bayes estimators provided can be readily used without requiring further derivations. The prior distribution used here is simpler than the one used in AL-Hussaini and Jaheen (1994) but one has to take caution while applying this class of prior as it assumes the independence of the two parameters.

We have shown an application of our methods to a carcinoma survival data set. Using the data, we have shown that the Bayes estimators based on the proposed prior is insensitive to the choice of the hyperparameters.

6 Acknowledgements

We acknowledge the remarks made by an anonymous referee which led to the improvement of the manuscript.

A Lindley's Approximation

For an easy reference for the readers, we describe Lindley's approximation method here for two parameters. The Bayes estimator of a function $u(\alpha,\beta)$ involves the evaluation of ratios of integrals of the form:

$$I(\mathbf{x}) = \frac{\int u(\alpha,\beta) e^{L(\alpha,\beta|\mathbf{X})+\rho(\alpha,\beta)} d(\alpha,\beta)}{\int e^{L(\alpha,\beta|\mathbf{X})+\rho(\alpha,\beta)} d(\alpha,\beta)}, \tag{A.1}$$

where $L(\alpha,\beta|\mathbf{x}) = \ln \ell(\alpha,\beta|\mathbf{x})$ is the log-likelihood and $\rho(\alpha,\beta) = \ln \pi(\alpha,\beta)$ is the log-prior. Let $(\hat{\alpha},\hat{\beta})$ denote the MLE of (α,β), *i.e.,* ,

$$(\hat{\alpha},\hat{\beta}) = \arg\max_{(\alpha,\beta)} \ell(\alpha,\beta|\mathbf{x}). \tag{A.2}$$

According to Lindley (1980), under certain regularity conditions, the ratio of integrals $I(\mathbf{x})$ can be approximated by

$$\begin{aligned} I(\mathbf{x}) \quad &\doteq \quad u(\hat{\alpha},\hat{\beta}) + \frac{1}{2}\{(\hat{u}_{\alpha\alpha} + 2\hat{u}_\alpha\hat{\rho}_\alpha)\,\hat{\sigma}_{\alpha\alpha} + (\hat{u}_{\beta\alpha} + 2\hat{u}_\beta\hat{\rho}_\alpha)\,\hat{\sigma}_{\beta\alpha} + (\hat{u}_{\alpha\beta} + 2\hat{u}_\alpha\hat{\rho}_\beta)\,\hat{\sigma}_{\alpha\beta} + \\ &\quad (\hat{u}_{\beta\beta} + 2\hat{u}_\beta\hat{\rho}_\beta)\,\hat{\sigma}_{\beta\beta} + (\hat{u}_\alpha\hat{\sigma}_{\alpha\alpha} + \hat{u}_\beta\hat{\sigma}_{\alpha\beta})\left(\hat{L}_{\alpha\alpha\alpha}\hat{\sigma}_{\alpha\alpha} + \hat{L}_{\alpha\beta\alpha}\hat{\sigma}_{\alpha\beta} + \hat{L}_{\beta\alpha\alpha}\hat{\sigma}_{\beta\alpha} + \hat{L}_{\beta\beta\alpha}\hat{\sigma}_{\beta\beta}\right) \\ &\quad + (\hat{u}_\alpha\hat{\sigma}_{\beta\alpha} + \hat{u}_\beta\hat{\sigma}_{\beta\beta})\left(\hat{L}_{\alpha\alpha\beta}\hat{\sigma}_{\alpha\alpha} + \hat{L}_{\alpha\beta\beta}\hat{\sigma}_{\alpha\beta} + \hat{L}_{\beta\alpha\beta}\hat{\sigma}_{\beta\alpha} + \hat{L}_{\beta\beta\beta}\hat{\sigma}_{\beta\beta}\right)\Big\}, \end{aligned} \tag{A.3}$$

where, for instance, $u_{\alpha\alpha}$ denotes the second derivative of the function $u(\alpha,\beta)$ with respect to α and $\hat{u}_{\alpha\alpha}$ represents the same expression evaluated at $\alpha = \hat{\alpha}$ and $\beta = \hat{\beta}$. The σ-terms are defined as $\sigma_{\alpha\alpha} = -\frac{1}{L_{\alpha\alpha}}$, $\sigma_{\beta\beta} = -\frac{1}{L_{\beta\beta}}$.

B Tierney and Kadane's Approximation

Suppose $u(\alpha,\beta)$ is the parametric function of interest. Define $l(\alpha,\beta) = n^{-1}\{L(\alpha,\beta;x) + \rho(\alpha,\beta)\}$, where ρ is the log-prior and L is the likelihood. In addition, define $l^*(\alpha,\beta) =$

$n^{-1} \ln u(\alpha, \beta) + l(\alpha, \beta)$. Further let $(\hat{\alpha}_l, \hat{\beta}_l)$ and $(\hat{\alpha}_{l^*}, \hat{\beta}_{l^*})$ be the values of (α, β) at which the functions l and l^* respectively attains their maximum. Then the integral

$$I(\mathbf{x}) = \int e^{nl^*(\alpha,\beta)} d(\alpha, \beta) \Big/ \int e^{nl(\alpha,\beta)} d(\alpha, \beta)$$

can be approximated by

$$\hat{I}(\mathbf{x}) = \sqrt{\frac{|\Sigma^*|}{|\Sigma|}} exp \left[n \left\{ l * (\hat{\alpha}_{l^*}, \hat{\beta}_{l^*}) - l(\hat{\alpha}_l, \hat{\beta}_l) \right\} \right], \tag{B.1}$$

where $|\Sigma|$ and $|\Sigma^*|$ are the negatives of inverse Hessians of $l(\alpha, \beta)$ and $l^*(\alpha, \beta)$ respectively evaluated at $(\hat{\alpha}_l, \hat{\beta}_l)$ and $(\hat{\alpha}_{l^*}, \hat{\beta}_{l^*})$.

References

[1] AL-Hussaini, E. K. and Jaheen, Z. F. (1994). Approximate bayes estimators applied to the burr model. *Communications in Statistics - Theory and Methods* **23 (1)**, 99–121.

[2] Burr, I. W. (1942). Cumulative frequency functions. *Annuls of Mathematical Statistics* **13**.

[3] Kalbfleisch, J. D. and Prentice, R. L. (2003). *The Statistical Analysis of Failure Time Data*. Wiley-Interscience, New York.

[4] Lindley, D. V. (1980). Approximate Bayesian methods. In Bernardo, J. M., De Groot, M. H., Lindley, D. V. and Smith, A. F. M., editors, *Bayesian Statistics*, pages 223–245, Spain. Valencia Press.

[5] Papadopoulos, A. S. (1978). The Burr distribution as a failure model from a bayesian approach. *IEEE Transactions on Reliability* **27**.

[6] Press, S. J. (2003). *Subjective and Objective Bayesian Statistics: Principles, Models and Applications*. Wiley, New York.

[7] Rodriguez, R. N. (1977). A guide to the Burr type XII distributions. *Biometrika* **64**, 129–134.

[8] Soliman, A. A. (2002). Reliability estimation in a generalized life-model with application to the burr-xii. *IEEE Transactions on Reliability* **51 (3)**, 337–34a.

[9] Soliman, A. A. (2005). Estimation of parameters of life from progressively censored data using burr-xii model. *IEEE Transactions on Reliability* **54 (1)**, 34–42.

[10] Varian, H. R. (1975). A Bayesian approach to real estate assessment. In Fienberg, S. E. and Zellner, A., editors, *Studies in Bayesian Econometrics and Statistics in honor of L. J. Savage*, pages 195–208, Amsterdam. North Holland Pub. Co.

[11] Wingo, D. R. (1983). Maximum likelihood methods for fitting the burr type xii distribution to life test data. *Biometrical Journal* **25**, 77–84.

[12] Wingo, D. R. (1993). Maximum likelihood methods for fitting the burr type xii distribution to multiply (progressively) censored life test data. *Metrika* **40**, 203–210.