RESEARCH SUPPORT
UNIVERSITY OF MICHIGAN BUSINESS SCHOOL

DECEMBER 1996

# A Bayesian Bivariate Failure Time Regression Model

## Working Paper #9612-35

PAUL DAMIEN
UNIVERSITY OF MICHIGAN BUSINESS SCHOOL

AND

PETER MULLER
ISDS, DUKE UNIVERSITY

# A BAYESIAN BIVARIATE FAILURE TIME REGRESSION MODEL

By PAUL DAMIEN

School of Business Administration

University of Michigan, Ann Arbor. MI 48109-1217. USA

email: damien@damien.bus.umich.edu

and PETER MÜLLER

ISDS. Duke University. Durham NC 27708-0251. USA

(919) 684 3437

email: pm@stat.duke.edu

# A BAYESIAN BIVARIATE FAILURE TIME REGRESSION MODEL

## SUMMARY

In this paper. we model bivariate failure time data under the assumption that the two random quantities corresponding to each of the failure times are dependent. We extend the analysis from a bivariate Gumbel model by regressing the log marginal hazard rate against context-motivated covariates. Illustrative analyses using data from electrical treeing and cancer relapse times are exemplified using Markov Chain Monte Carlo methods.

*Keywords:* Gumbel Distribution, Gibbs Sampling, Correlation. Failure Time, Hazard Rate. Covariates.

# 1 INTRODUCTION

We use the term failure time to refer both to reliability data arising in engineering studies and survival data arising in medical or clinical trials. In engineering studies, reliability is defined as the probability that the system has not failed at time $t$. Likewise, in clinical trials, survival is defined as the probability that a patient has not died, or contracted the disease under investigation at time $t$. Most work in this area has been concerned with the case of one observed failure time while in the case of two observed failure times, most authors, using a frequentist approach, have considered such failure times to be independent: see, for example, Mann and Grubbs[1]; Guttman and Sinha[2]; Miyamura[3].

Tawn[4] considers bivariate models and estimation by developing an asymptotic estimate for the dependence function. Tawn[5] extends the bivariate structure to multivariate models using extreme value distributions. Oakes[6] also considers a non-Bayesian analysis of bivariate models that are induced by frailties. The Gumbel model is studied in that paper. Most recently, Oakes and Manatunga[7] obtain an asymptotic estimate for the Fisher information for a bivariate extreme value distribution. Thus, there appears to be several useful papers in recent literature that seek to provide asymptotic estimates for the parameters of bivariate failure time distributions. Recent Bayesian work in this area is, however, somewhat sparse, the notable exception being Draper et al.[8].

Our approach extends the work of Draper et al.[8], who provide a Bayesian analysis of system reliability when two components are dependent, in two respects. Firstly, we bypass the need for numerical integration of awkard posterior forms via Monte Carlo methods that are straightforward to implement. In addition to ease in implementation, these methods provide a full Bayesian solution to the problem that Draper et al.'s method does not accomplish. Secondly, we extend the bivariate Gumbel model by modelling the logarithm of the marginal hazard rates through a regression set-up. A full Bayesian analysis for the latter model is provided and exemplified using bladder cancer data.

## 2 THE BIVARIATE GUMBEL MODEL

Suppose two components have failure times $(T_1, T_2)$ that are distributed with density function

$$f(t_1, t_2) = \lambda_1 \lambda_2 [exp(-\lambda_1 t_1 - \lambda_2 t_2)][1 + \alpha h_1(t_1) h_2(t_2)]. \tag{1}$$

where $\lambda_1, \lambda_2, t_1, t_2 > 0, -1 \leq \alpha \leq 1$. and $h_j$ is defined by

$$h_j(t_j) = 2exp(-\lambda_j t_j) - 1. \quad j = 1, 2.$$

This is the Gumbel bivariate exponential distribution (Gumbel[9]). The marginals of this joint density are

$$f_j(t_j) = \lambda_j exp(-\lambda_j t_j).$$

with $E(T_j) = \lambda_j^{-1}$ and $\mathrm{Var}(T_j) = \lambda_j^{-2}$. The reliability of a system using the model in (1) is given by

$$R(t) = P(T_1 > t, T_2 > t) = exp(-\lambda t)[1 + \alpha \{e^{-\lambda_1 t} - 1\}\{e^{-\lambda_2 t} - 1\}]. \tag{2}$$

with $\lambda = \lambda_1 + \lambda_2$.

Moeschberger and Klein[10] have argued, from a frequentist point of view. that the consequences of assuming $\alpha = 0$ (independence) in the above set-up could be very misleading. Draper et al. using a Bayesian approach, do not assume that $\alpha = 0$. but, they are only able to provide estimates of the posterior mean of $R(t)$ by numerically integrating out appropriate random variables after a scaling transformation of the Bayesian model described below.

## 3 THE BAYESIAN MODEL

Let $t_i = (t_{1i}, t_{2i})'$. $i = 1, \ldots, n$ be $n$ independent observations on $(T_1, T_2)$ whose density is given by (1), so that denoting $(t_1, \ldots t_n)$ by $T$, the likelihood function is given by

$$f(T|\lambda_1, \lambda_2, \alpha) \propto \lambda_1^n \lambda_2^n exp(-(\lambda_1 S_1 + \lambda_2 S_2)) \prod_{i=1}^{n} [1 + \alpha h_1(t_{1i}) h_2(t_{2i})]. \tag{3}$$

with $S_j = \sum_{i=1}^{n} t_{ji}, j = 1, 2$. Denoting the prior joint distribution for $(\lambda_1, \lambda_2, \alpha)$ by $[\lambda_1, \lambda_2, \alpha]$, the posterior joint distribution of $(\lambda_1, \lambda_2, \alpha)$ is given, by Bayes' Theorem. as

$$[\lambda_1, \lambda_2, \alpha|T] \propto f(T|\lambda_1, \lambda_2, \alpha)[\lambda_1, \lambda_2, \alpha], \tag{4}$$

4

where the notation $[.]$ denotes the probability density of $\cdot.\cdot$. In this paper. following Draper et al... we will only consider priors of the form

$$[\lambda_1. \lambda_2. \alpha] \propto (1 - \alpha)^{r_1-1}(1 + \alpha)^{r_2-1} \prod_{j=1}^{2} \lambda_j^{p_j-1} exp(-\lambda_j q_j). \qquad (5)$$

Equation (5) implies that $(1 - \alpha)/2$ has a Beta $(r_1. r_2)$ distribution. $\lambda_j$ has a Gamma $(p_j. q_j)$ distribution. and that the three parameters are mutually independent. Under certain conditions. (5) can be shown to reduce to the non-informative prior

$$[\lambda_1. \lambda_2. \alpha] \propto \lambda_1^{-1}\lambda_2^{-1}, \quad \lambda_1, \lambda_2 > 0, \; -1 \leq \alpha \leq 1. \qquad (6)$$

Draper et al. use numerical integration methods after suitably transforming (4) to obtain the exact posterior distribution of the three parameters.

## 4 Bivariate Gumbel exponential regression

We note that our choice of the bivariate gumbel distribution in this paper is purely to illustrate the Bayesian method when one observes two random quantities (failure times) that may be correlated. Other bivariate models. such as the Pareto. exist and may be more appropriate in a given context.

Often it is the case that failure times in clinical trials are related to covariates such as gender. age etc. We include a regression on covariates by extending the model in (3) to

$$f(t_{1i}. t_{2i}|x, \lambda_1^0, \lambda_2^0, \alpha, \beta_1. \beta_2) = \lambda_1. \lambda_2[exp(-\lambda_1 t_1 - \lambda_2 t_2)][1 + \alpha h_1(t_1, \lambda_1)h_2(t_2, \lambda_2)].$$
$$\log(\lambda_1) = \log \lambda_1^0 + \beta_{1,1}x_{1i} + \ldots + \beta_{1,m}x_{mi},$$
$$\log(\lambda_2) = \log \lambda_2^0 + \beta_{2,1}x_{1i} + \ldots + \beta_{2,m}x_{mi}, \qquad (7)$$

where $x = (x_{1i}, x_{2i})$ $i = 1, \ldots m$ is the vector of $m$ observations on each covariate for each of the two failure times. $t_1$ and $t_2$. We complete the model by specifying a prior distribution on the $\beta$ vectors taking these to have a multivariate normal distribution.

$$\beta_j \sim \mathcal{N}(b. B). \; j = 1, 2.$$

We assume that $\beta$ and $(\lambda_1, \lambda_2. \alpha)$ are a priori independent. Note that any other prior distribution on $\beta$ is possible, and our choice of the normal distribution is simply to exemplify the Bayesian approach to the problem.

## 5   COMPUTATIONS

By a combination of Metropolis and independence chain steps we build a Markov chain Monte Carlo scheme to estimate the Gumbel bivariate exponential model. For a review of Markov chain Monte Carlo methods see. for example. Smith and Roberts[11]. Tierney[12] or Gilks et al.[13]. The underlying rationale of Markov chain Monte Carlo methods is to simulate a Markov chain which is constructed to have the desired posterior as the limiting distribution. Under conditions discussed in Tierney[12] appropriate ergodic averages over the simulated chain can then be used to estimate posterior expectations. marginal posterior distributions, etc.

### 5.1   Estimating the Gumbel bivariate exponential model

We discuss first the Markov chain used for the estimation of the bivariate exponential model in (3) with the prior given in (5). Starting with initial values for the three parameters $\lambda_1, \lambda_2, \alpha$ we define a chain by the following three steps.

First, draw a "candidate" $\tilde{\lambda}_1 \sim Gamma(a, b)$ with $a = p_1 + n$ and $b = q_1 + \sum t_{1i}$. Denote by $g(\lambda)$ the Gamma(a,b) p.d.f. Compute

$$a(\lambda_1, \tilde{\lambda}_1) = \min\left(1, \frac{p(\tilde{\lambda}_1|\lambda_1, \alpha, T)\, g(\lambda_1)}{p(\lambda_1|\lambda_1, \alpha, T)\, g(\tilde{\lambda}_1)}\right) = \frac{\prod_{i=1}^{n}[1 + \alpha h_1(t_{1i}, \tilde{\lambda}_1) h_2(t_{2i}, \lambda_2)]}{\prod_{i=1}^{n}[1 + \alpha h_1(t_{1i}, \lambda_1) h_2(t_{2i}, \lambda_2)]}.$$

With probability $a$ set $\lambda_1 := \tilde{\lambda}_1$. else leave the current value of $\lambda_1$ unchanged. This is an implementation of an independence chain step, using the Gamma posterior of the marginal model in $t_{1i}$ as approximation of the full conditional posterior. See Tierney[12] for more discussion.

Second. replace $\lambda_2$ by repeating the independence chain desrcibed above for $\lambda_2$.

Third. update $\alpha$ by implementing a Metropolis step: Generate $\tilde{\alpha} \sim N(\alpha, \sigma_\alpha)$ and evaluate

$$a(\alpha, \tilde{\alpha}) = \min\left(1, \frac{p(\tilde{\alpha}|\lambda_1, \lambda_2, T)}{p(\alpha|\lambda_1, \lambda_2, T)}\right).$$

With probability $a$ replace $\alpha$ by $\tilde{\alpha}$. else keep the currently imputed value of $\alpha$: see Section 6.3 for comments on choosing the arbitrary scale factor $\sigma_\alpha$.

## 5.2 Extension to the Gumbel bivariate regression model

The Markov chain Monte Carlo scheme described for model (1) is easily modified to include the estimation of regression parameters in the Gumbel bivariate regression model (7).

Starting with initial values for the paramter vector $\theta = (\lambda_1^0, \lambda_2^0, \alpha, \beta_{11} \ldots \beta_{2m})$. we simulate a Markov chain defined by iteratively scanning over all parameters. replacing each parameter by simulating a Metropolis step. Relabel the parameter vector as $\theta = (\theta_1, \ldots, \theta_p)$ (i.e. write $\theta_1$ for $\lambda_1^0$ etc.). Each Metropolis step is given by: First draw a "candidate" $\tilde{\theta}_j \sim N(\theta_j, \sigma_j)$ and compute

$$a(\theta_j, \tilde{\theta}_j) = \min\left(1, \frac{p(\tilde{\theta}_j | \theta_{-j}, T)}{p(\theta_j | \theta_{-j}, T)}\right).$$

where $\theta_{-j}$ denotes the full parameter vector without $\theta_j$. Second. with probability $a$ replace $\theta_j$ by $\tilde{\theta}_j$, else keep the currently imputed value of $\theta_j$. The $\sigma_j$ are arbitrary scale factors. Section 6.3 suggests some default choices. Note this is identical to the Metropolis step which was used in simulating the $\alpha$ parameter in section 5.1.

## 6 EXAMPLES

### 6.1 Example 1: A Gumbel bivariate exponential

To illustrate the proposed Markov chain Monte Carlo scheme we estimated the model with the electrical cable failure data from Lawless[14]. The same data set was used by Draper et al.[8]. The data concerns time to inception of a microscopic defect in the material $(t_{1i})$ and the subsequent additional time to eventual failure $(t_{2i})$. This phenomenon is commonly referred to as electrical treeing in the engineering literature.

The posterior marginal distributions of the three parameters are plotted in Figures 1 and 2. In addition. easily obtained forms of inference summaries using the Monte Carlo samples provide a complete description of the uncertainty in the three parameters. It is interesting to observe that there is considerable evidence from the data that $\alpha$ is non-zero, suggesting a strong dependence between the two failure times. We note that the results from this simulation are in close agreement with results obtained using exact numerical (quadrature) methods by Draper et al.[8]. The key function of interest is the reliability

function given in (2). A full description of the uncertainty in the reliability function at each observed time is easily obtained by simply evaluating (2) at each of the simulated samples. As an illustration, a complete description of the uncertainty in $R(t)$ at times $t = 50, 100, 150$ is provided in Figure 3.

We note, once again, that Draper et al. only provide an estimate of the mean of $R(t)$ at each point along the time axis.

## 6.2 Example 2: A Gumbel bivariate exponential regression model

Byar et al.[15] report a study comparing placebo, pyridoxine and topical thiotepa in preventing recurrence of Stage I bladder cancer. Stage I tumors tend to recur in around 50% of the patients within two years after removal. Instillations of thiotepa and regular ingestion of pyridoxine are believed to retard the growth of tumors, or cure them altogether.

We consider patients with 2 or more recurrence times. For each of the $n = 28$ patients we record the two relapse times $(t_{1i}, t_{2i}$, two indicators for the assigned treatment $(x_{1i} = 1$ for pyridoxine, $x_{2i} = 1$ for thiotepa, $x_{1i} = x_{2i} = 0$ for placebo), and number $(x_{3i})$ and size $(x_{4i})$ of initial bladder cancer occurences. We estimate the bivariate life time regression model (7) assuming a noninformative prior (6) for $(\lambda_1, \lambda_2, \alpha)$, and taking

$$f(t_{1i}, t_{2i} | x_i, \lambda_1^0, \lambda_2^0, \alpha, \beta_1, \beta_2) = \lambda_1 \lambda_2 [\exp(-\lambda_1 x_1 - \lambda_2 x_2)][1 + \alpha h_1(x_1, \lambda_1)h_2(x_2, \lambda_2)].$$

$$\log(\lambda_1) = \lambda_1^0 + \beta_{1,1}x_{1i} + \beta_{1,2}x_{2i} + \beta_{1,3}x_{3i} + \beta_{1,4}x_{4i},$$

$$\log(\lambda_2) = \lambda_2^0 + \beta_{2,1}x_{1i} + \beta_{2,2}x_{2i} + \beta_{2,3}x_{3i} + \beta_{2,4}x_{4i}.$$

A strength of the Bayesian approach is to enable the statistician to factor in context-dependent constraints. Thus, a priori, we believe, based on medical knowledge, that the coefficients in the $\beta$ vectors corresponding to the covariates, number, and size of initial bladder cancer occurences are non-negative. We factor this qualitative information by restricting the appropriate components in the $\beta$ vectors in the prior distribution to be non-negative.

Figure 4 shows the estimated marginal posterior distributions for $\lambda_1$, $\lambda_2$ and $\alpha$. Since inference is based on simulated Monte Carlo samples, any desired inference summaries can be provided. Like in the previous example, we observe strong dependency between the two failure times as depicted in Panel (b) of Figure 4. As an illustration, a complete posterior description for $\beta_{j1}$ and $\beta_{j2}$ in the regression model is given in Figure 5. The reliability function in (2) in the

context of these data is the cumulative survival based on the two recurrence times. As an illustration. a plot of the survival function is provided in Figure 6a.

Again, as an illustration. in Figures 7a and b. we plot the posterior marginal distribution in the form of histograms for the coefficients $\beta_{13}$ and $\beta_{14}$, respectively. Recall that these coefficients correspond to the effects of prevalance ($x_3$) and size ($x_4$), respectively. of initial bladder occurence. Moreover. as noted previously, the prior distribution incorporates the restriction that both of these parameters are non-negative.

A remarkable feature of Monte Carlo based methods in Bayesian inference is that questions of practical interest such as the following can be answered straightforwardly: "what is the posterior probability that a patient will have first and second relapse times greater than 5 months if the patient is subjected to thiotepa treatment and intial tumor size and prevalance are $x_{3i} = 1$ and $x_{4i} = 1$?" Using the simulated samples for each of the parameters, we simply calculate, for each sample, the required probability from (2), thus obtaining a complete description of the uncertainty for this prediction. This is reported in Figure 6b.

## 6.3  Implementation and initialization

The Markov chain Monte Carlo scheme proposed in Section 5.2 implements a Metropolis-within-Gibbs algorithm as proposed in Müller[17]. In the examples reported in this paper we chose the scale factors $\sigma_j$ by approximately estimating posterior standard deviations via Laplace integration first. For Example 2. we estimated a simplified univariate regression model in $t_{i1}$ only, using the approximate posterior standard deviations of $\lambda_1^0, \beta_{1,j}$ for both, the scale factors on $\lambda_1^0$ and $\beta_{1,j}$ as well as $\lambda_2^0$ and $\beta_{2,j}$. Based on the restricted parameter space for $\alpha$ we set $\sigma_\alpha = 0.5$. The approximate posterior means from the Laplace integration served as initial values to start the Markov chain Monte Carlo simulations. We, arbitrarily. initialized $\alpha$ at $\alpha = 0$.

Another issue to be resolved for practical implementation of the scheme is how to decide termination of the simulation scheme. In the reported examples we started by running 5000 iterations of the chain. Then we evaluated the convergence diagnostic proposed by Geweke[17]. In Example 1, the diagnostic indicated that the chain had already practically converged. For Example 2, the diagnostic indicated convergence after 15,000 iterations.

9

Both Markov chain Monte Carlo schemes were implemented as C programs on a Digital workstation 3500 check. The 5000 iterations in Example 1 required 27 seconds CPU time. The 15,000 iterations in Example 2 took 23 minutes CPU time.

To plot posterior marginal distributions like in Figure 1 we used a "Rao-Blackwellization" type estimate as proposed in Gelfand and Smith[18]. For example, let $\theta^{(t)}$ denote the imputed parameter vector after $t$ iterations of the Markov chain. To estimate $p(\lambda_1|T)$ for $\lambda_1 = 0.002$ we evaluate $p(\lambda_1|\lambda_2^{(t)}, \alpha^{(t)}, T)$ for $t = 100, 150, 200, \ldots, 5000$ and compute $\hat{p}(\lambda_1|T) = \frac{1}{M}\sum p(\lambda_1|\lambda_2^{(t)}, \alpha^{(t)}, T)$. We only use every 50-th iteration for reasons of computational efficiency.

## 7 SOME EXTENSIONS

In this paper, we have considered modelling two failure times by allowing for potential dependence in the random quantities corresponding to each failure time. Several extensions are possible. First, other bivariate models, such as the bivariate Pareto distribution, could be easily considered using a similar framework. Secondly, the implications of such models in competing risk analysis needs investigation. Thirdly, hierarchical models and nonparametric generalisations for the marginal hazard rate are possible. For example, one could model the hyperparameters of the regression model as realisations of a stochastic process; see, for example, West, Mueller & Escobar[19]. Fourthly, following Tawn[5], the computational details for a multivariate extension from a Bayesian perspective needs study. These and other extensions will be reported elsewhere.

## REFERENCES

1. Mann, N.R. & Grubbs, F.E., 'Approximately optimum confidence bounds for system reliability based on component test data', *Technometrics*, **61**, 335-347 (1974).

2. Guttman, J. & Sinha, S.K.. 'Bayesian inference about the reliability function for the exponential distributions', *Communications in Statistics. A.5*, 471-479 (1976).

3. Miyamura, T., 'Estimating component failure rates from combined component and systems data: Exponentially distributed component lifetimes,' *Technometrics*, **24**, 313-318 (1982).

4. Tawn, J.A., 'Bivariate extreme value theory: models and estimation'. *Biometrika*, **75**, 397-415 (1988).

5. Tawn, J.A., 'Modelling multivariate extreme value distributions', *Biometrika*, **77**, 245-253 (1990).

6. Oakes, D., 'Bivariate survival models induced by frailties', *Journal of the American Statistical Association*, **84**, 487-493 (1989).

7. Oakes, D. & Manatunga, A.K., 'Fisher information for a bivariate extreme value distribution', *Biometrika*, **79**, (1992).

8. Draper, N.R., Evans, M., & Guttman, I., 'A Bayesian approach to system reliability when two components are dependent', *Computational Statistics and Data Analysis* , **7**, 39–49 (1977).

9. Gumbel, E.J., 'Bivariate exponential distributions', *Journal of the American Statistical Association*, **55**, 698-707, (1960).

10. Moeschberger, M.L. & Klein, J.P., 'Consequences of departures from independence in exponential series system', *Technometrics*, **26**, 277-284, (1984).

11. Smith, A.F.M. & Roberts, G.O., 'Bayesian computation via the Gibbs sampler and related Markov chain Monte Carlo methods (with discussion)', *Journal of the Royal Statistical Society, Series B*, **55**, 3-23, (1993).

12. Tierney, L., 'Markov chains for exploring posterior distributions', *Annals of Statistics*, **22**, 1701-1728 (1994).

13. Gilks, W.R., Clayton, D.G., Spiegelhalter, D.J., Best, N.G., McNeil, A.J., Sharples, L.D., & Kirby, A.J., 'Modelling complexity: applications

of Gibbs sampling in medicine'. *Journal of the Royal Statistical Society B.* **55**. 39-52 (1993).

14. Lawless, J.F.. *Statistical Models and Methods for Lifetime Data.* New York: Wiley. 1982.

15. Byar. D.P.. Blackard. C. & VACURG 'Comparison of Placebo. Pyridoxine and Topical Thiotepa in Preventing Recurrence of Stage I Bladder Cancer', *Urology,* 10, pp 556-561 (1977).

16. Mueller, P.. 'Metropolis Based Posterior Integration Schemes'. Technical Report 91-09. Statistics Department, Purdue University. (1991).

17. Geweke, J.. 'Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments', in *Bayesian Statistics 4*, ed. J.O. Berger, J.M. Bernardo, A.P. Dawid and Smith, A.F.M., 169-194 (1992).

18. Gelfand, A.E. and Smith A.F.M. 'Sampling based approaches to calculating marginal densities', *Journal of the American Statistical Association.* **85**, 398-409 (1990).

19. West, M., Mueller, P., & Escobar, M.D., 'Hierarchical priors and mixture models, with application in regression and density estimation', in *Aspects of Uncertainty: A tribute to D. V. Lindley,* ed. A.F.M. Smith and P. Freeman, Wiley, New-York. (1994).

12
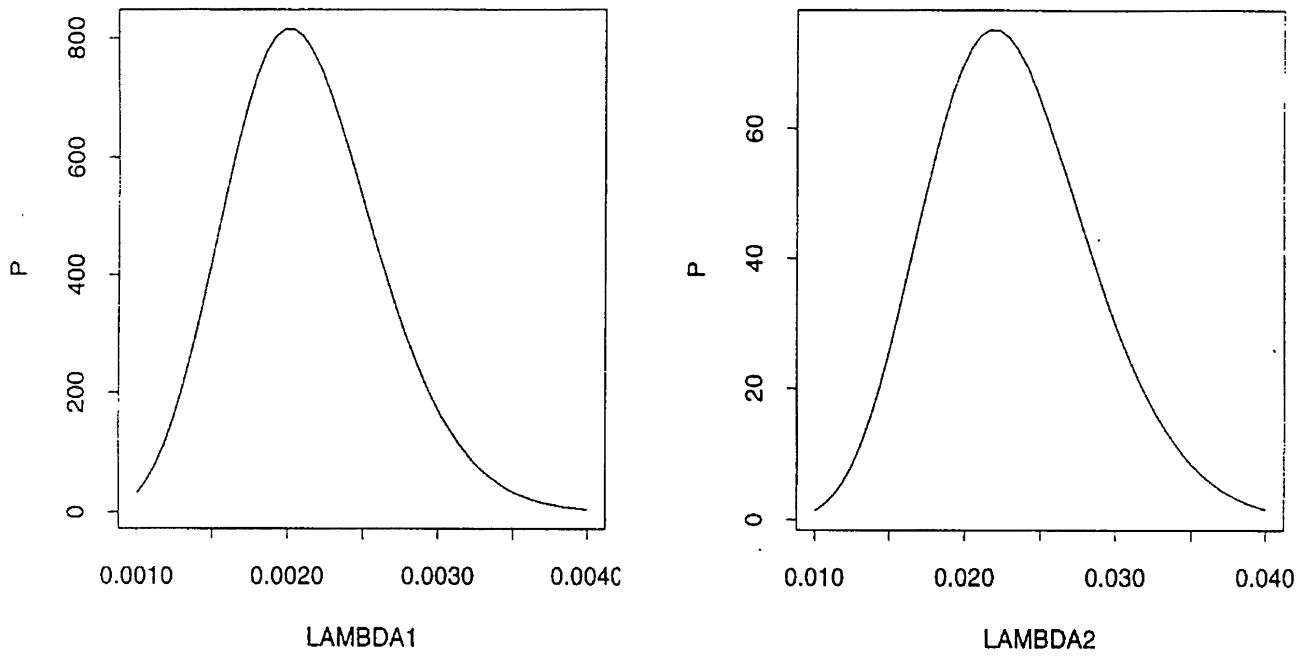
Figure 1: Example 1. Marginal posterior distributions $p(\lambda_1|x)$ and $p(\lambda_2|x)$. First quartile, median, mean and third quartile are estimated by $(Q_1 = 0.0018. Mdn = 0.0021, \bar{\lambda}_1 = 0.0021, Q_3 = 0.0024)$ and $(Q_1 = 0.019, Mdn = 0.023. \bar{\lambda}_2 = 0.023, Q_3 = 0.026)$ respectively.
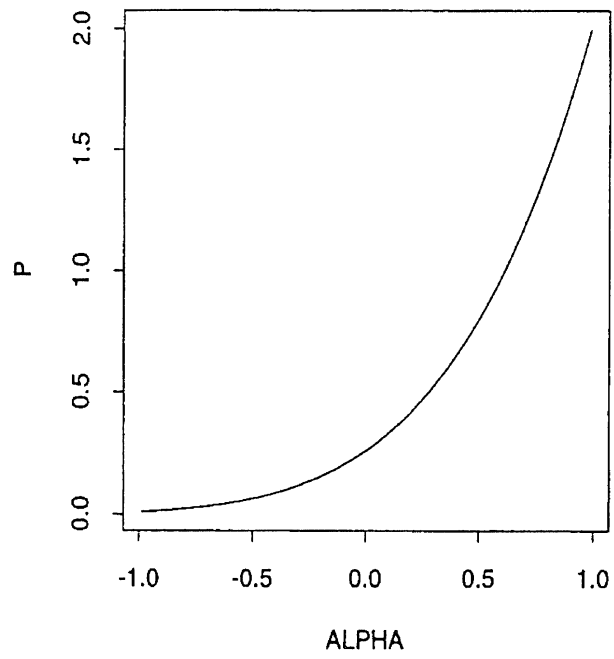
Figure 2: Example 1. Marginal posterior $p(\alpha|x)$ on the correlation parameter. Estimated posterior quartiles and mean are $Q_1 = 0.38$, $Mdn = 0.68$, $\bar{\alpha} = 0.58$ and $Q_3 = 0.86$.
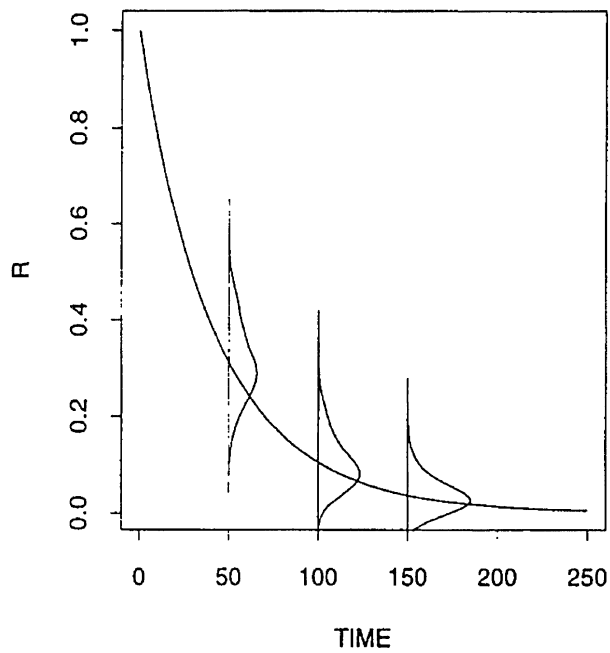
Figure 3: Example 1. The estimated reliability function $R(t) = P(X_1 > t, X_2 > t)$. The graph plots the posterior mean estimate for $R(t)$ against $t$. In addition to point estimates for $R(t)$ the Bayesian modelling framework provides a full description of the uncertainty about $R(t)$ at any $t$. For some selected values of $t$ the figure shows the full posterior distribution $p(R(t)|T)$.
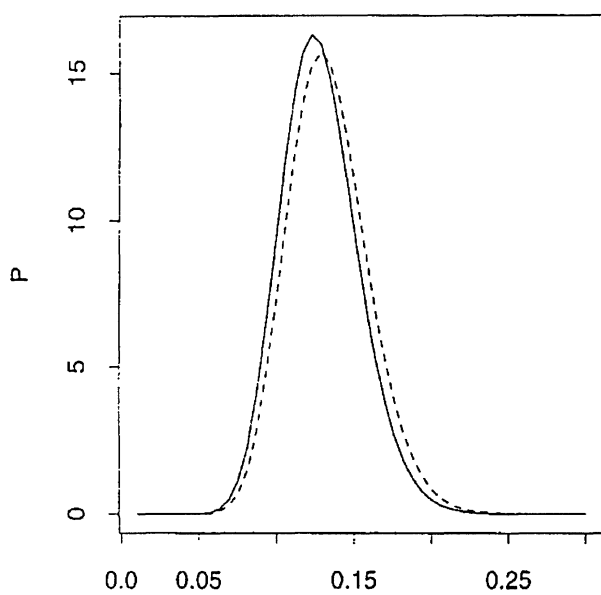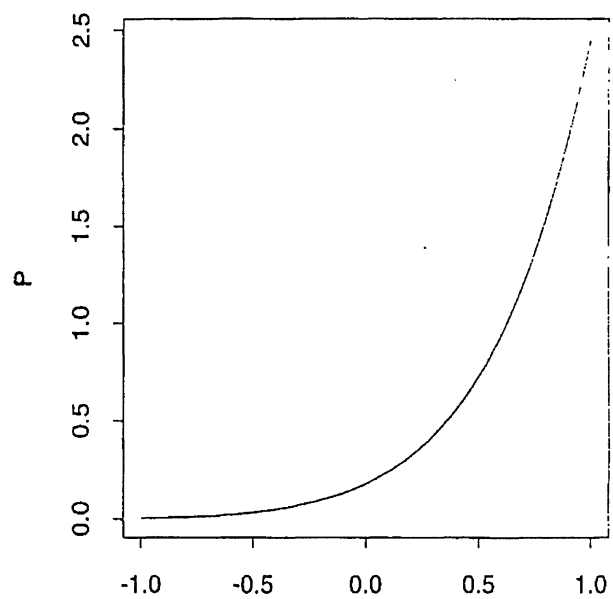
(a) $p(\lambda_j^0|x)$          (b) $p(\alpha|x)$.

Figure 4: Example 2. Panel (a) shows the posterior marginal of $\lambda_1^0$ (solid line) and $\lambda_2^0$ (dotted line). Panel (b) plots the posterior marginal of the correlation parameter $\alpha$.
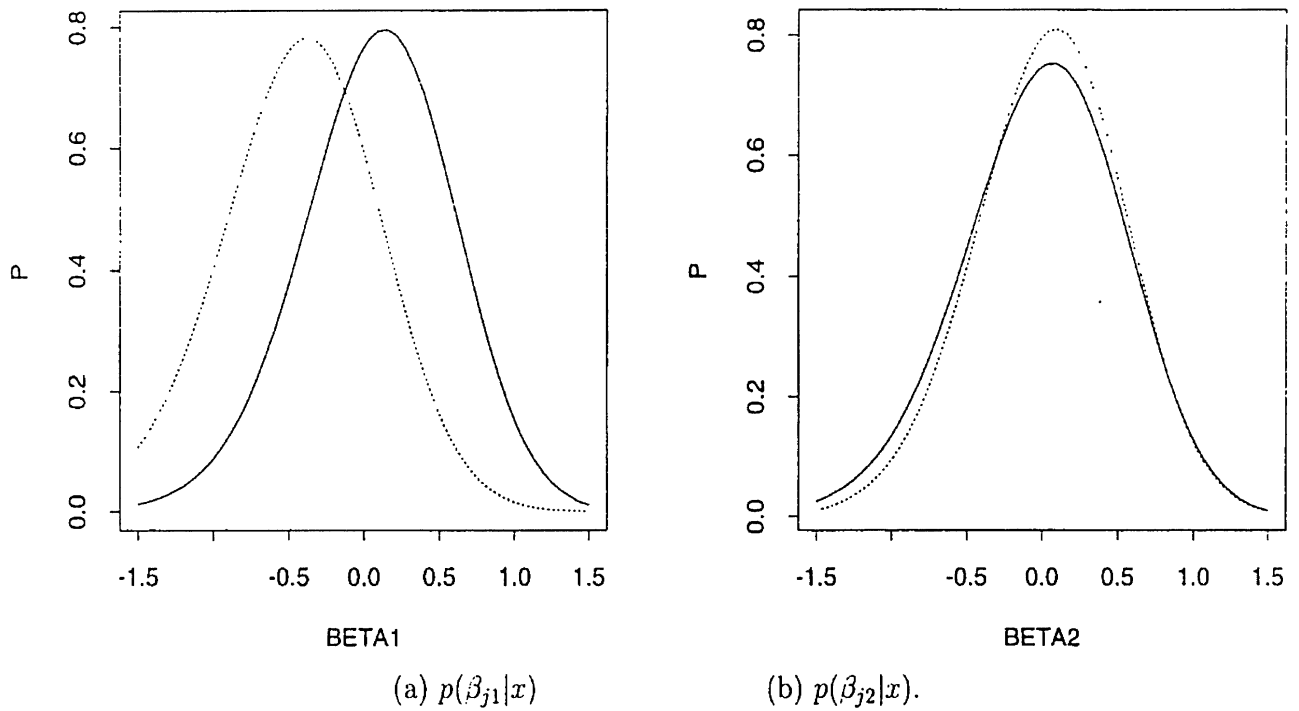
(a) $p(\beta_{j1}|x)$          (b) $p(\beta_{j2}|x)$.

Figure 5: Example 2. Posterior marginal distributions for the regression parameters $\beta_{11}$ (solid line in panel (a)), $\beta_{21}$ (dotted line in panel (a)) for first relapse time and $\beta_{12}$ (solid line in panel (b)) and $\beta_{22}$ (dotted line in panel (b)) for the second relapse time. The posterior distribution indicates a slightly lower log hazard for the first relapse time under the thiotepa treatment. For the second relapse time the two treatments are practically indistinguishable.
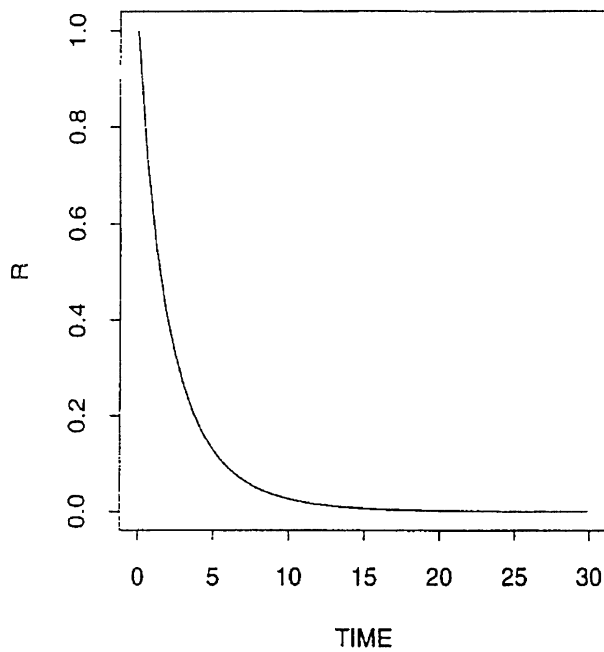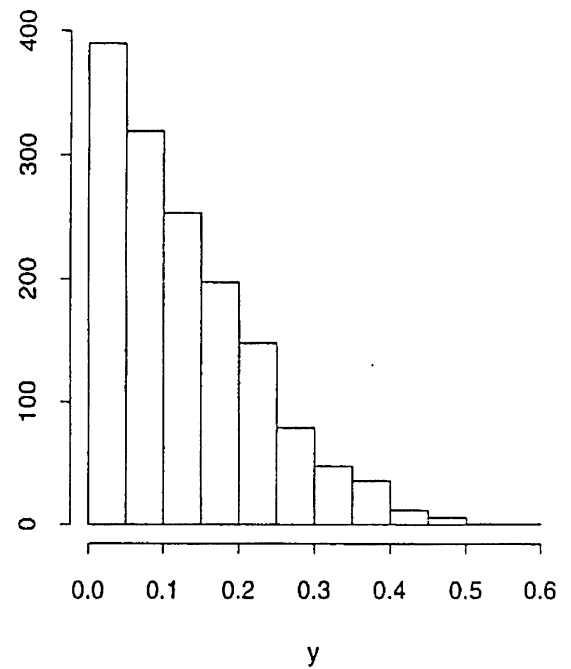
(a) $R(t)$                                   (b) $p(R(5)|x = (0,1,1,1), T)$

Figure 6: Example 2. The estimated reliability function $R(t) = P(T_1 > t. T_2 > t)$ for a patient with covariates $x = (0,1,1,1)$. In addition to point estimates for $R(t)$ the Bayesian modelling framework allows to put a posterior predictive distribution on quantities like the probability that a patient has first and second relapse time greater than 5 months when subjected to thiotepa treatment and with initial size and number of bladder cancer occurences equal $x_3 = 1$ and $x_4 = 4$. Panel (b) shows the histogram of the predictive posterior distribution for this probability.

18
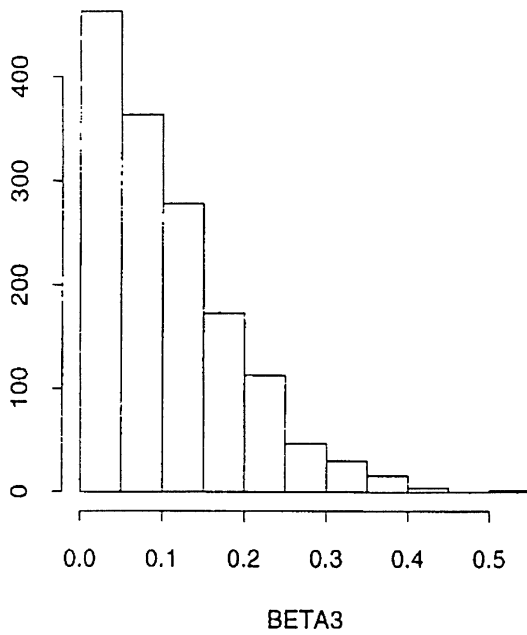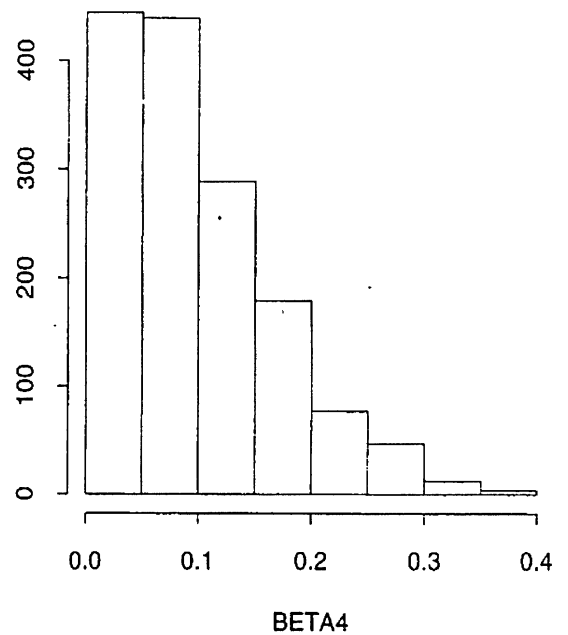
(a) $p(\beta_{13}|T)$              (b) $p(\beta_{14}|T)$.

Figure 7: Example 2. Posterior marginal distributions $p(\beta_{13}|T)$ (panel a), and $p(\beta_{14}|T)$ (panel b) for the the effects of prevalance ($x_3$) and size ($x_4$) of initial bladder cancer occurence. Here $T = (t_{1i}, t_{2i}, x_i, i = 1, \ldots, n)$ denotes the full data vector. The proposed Markov chain Monte Carlo scheme allows to include prior information on the parameters. In this example, we enforce *a priori* that $\beta_{13}$ and $\beta_{14}$ be non-negative.