Ayman A. Mostafa / International Journal of Engineering Research and Applications (IJERA)
ISSN: 2248-9622 www.ijera.com
Vol. 2, Issue 2,Mar-Apr 2012, pp.374-385

Semi-parametric hazard ratio applied to engineering insurance system

Ayman A. Mostafa
Department of Mathematics and Statistics, Al-Imam Muhammad Ibn Saud Islamic University, P.O. BOX 90950, 11623 Riyadh (SAUDI ARABIA)

Abstract
The objective of hazards (lifetime) analysis is to advance and promote statistical science in the various applied fields that deal with lifetime (survival) data including: actuarial science and reliability engineering. The lifetime data analysis provides special techniques that are required to compare the risks for failure. Bayesian semi-parametric methods have been applied to survival analysis problems since the emergence of the area of the Bayesian semi-parametric procedures. Cox proportional hazard model (PHM) estimates hazard ratios. Cox PHM is considered as constant hazard ratio over time if and only if Cox PHM assumptions are not violated. In Bayesian analysis, Markov Chin Monte Carlo (MCMC) methods have become a ubiquitous tool as the computer is more powerful. In this article, estimation of the parameters in Cox PHM is presented by using Bayes methods based on MCMC algorithm and duplicate the results using non-Bayes framework. The method is motivated by an example based on a hypothetical engineering insurance system.

Keywords: Cox proportional hazards model; posterior distributions; Markov chain Monte Carlo; WinBUGS.

MSC 2010 Mathematics Subject Classification: 62N99, 62F15, 62M20, 65C05, 65C40, 68N15.

1 Introduction
In the well known Cox proportional hazards model (PHM), introduced by Cox (1972) as a tool for analyzing lifetime data in the both of biomedical and engineering studies. It has become the standard nonparametric regression model for accelerated life testing in the past few years. The basic underlying assumption under the PHM is that the hazard rate is proportional to the effect of each explanatory (risk) variable. In other words, the effect of each explanatory variable is found by multiplying the baseline hazard by some function of the explanatory variable vector, which does not depend on time *t*. (see, for example, Cox and Oakes, 1984; Lane et al., 1986). The main advantages of the Cox proportional hazard model are that it explicitly incorporates the time to failure into the model and it requires no distributional assumption in estimating either the coefficients of the explanatory variables or baseline hazard function. Moreover, it is straightforward to estimate the parameters of interest in Cox PHM using common statistical packages, which is widely available these days. Meanwhile, one should consider two possible problems in employing the PHM. First, there might existing many ties among survival times for each observation (especially for the censored lifetime for non failures), which need to be corrected. Second, an assumption of time-constant explanatory variables has to be dealt with. More details will be explained in the later part of this section.

Hazard models, including the Cox proportional hazard model, have been successfully verified as an efficient classification and prediction tool for failure studies in the financial area (e.g., Kim et al., 1995; Lane et al., 1986; Whalen, 1991). The robustness of the model also has been proved in similar studies of other industries, in labor economics, and the social sciences (e.g., Chen and Lee, 1993; Kiefer, 1988; Ng, Cram and Jenkins, 1991).

The proportional hazards model has the following general form:

$$\lambda(t;Z) = \lambda(t \mid Z=0)\exp\left(\beta'Z\right) = \lambda_0(t)\exp(\beta'Z); \qquad (1)$$

This is a product of the unknown baseline hazard rate $\lambda_0(t)$ and the exponential function of the unknown regression coefficient. Typically; both β and Z are assumed to be constant over time t. The PHM is considered semi-parametric in that $\exp(\beta'Z)$ is parametric and the baseline hazard $\lambda_0(t)$ is nonparametric.

1.1 motivate the problem and derive model
In this paper, we will motivate the problem and derive model (1) as follows. Individual rating in non-life insurance may be based on many factors such as age of policy holder, urbanization,…,etc (Keiding, 1998). On the other hand, the reliability study for an engineering insurance system (EIS) is considered in this paper. Although it is a hypothetical example of a policy insurance contract, it could very well constitute a real situation. It is well known that any insurance business, it is critical to ensure the efficient management of insurance risk (see, www.fengineering.lt/en/solutions/insurance-risk). Consider an insurance company (insurer) to its owner' project (insured) a one year lifetime maintenance contract (insurance policy). This policy, guarantee to upgrade the product to replace with a new version without additional cost. Whenever, EIS failure occurs during the lifetime of the maintenance, the insured calls the insurance company to change this machine tool. Our insurance company is very interested in providing that system is reliable enough. To illustrate the analysis that will be presented in Section 3, we imagine that the EIS consists of the data in Table 1. The data come from Merrick and Soyer (2003).

The data consist of the failure times of $i = 1,\ldots,24$ machine tools and their corresponding explanatory variables are cutting speed (its label is $Z_{i,1}$), feed rate (its label is $Z_{i,2}$), and depth of cut (its label is $Z_{i,3}$). The 24 machine tools used for the cutting were tungsten carbide disposable inserts mounted in a tool holder. A 7.5-horsepower engine lathe equipped with a three-jaw universal chuck and a live center mounted in the tailstock was used to perform the cutting operation. The cutting operations were performed without using cutting fluids. It was pointed out in Merrick and Soyer (2003), the aim of machine tool life modeling is to aid decisions concerning the operation of machine tools. An important decision in the operation of machine tools is when to replace them.

The main goal of the insurance company is to estimate the hazard ratios in order to divide the risk into three parts:

1) Lower risk **2)** Intermediate risk **3)** Higher risk

This deviation of the risk helps the company to determine the suitable premiums and in the same time the insurance company achieves a reasonable profit. This study clearly constitutes as a real example of EIS. Reliability (survival function) is one of the key quality characteristics of components, products and systems. It cannot be directly measured and assessed like other quality characteristics but can only be predicted for given times and conditions. Its value depends on the use conditions of the product (risk factors) as well as the time at which it is to be predicted. Reliability prediction has a major impact on critical decisions such as the optimum release time of the product, the type and length of warranty policy and associated duration and cost, and the determination of the optimum maintenance and replacement schedules. Therefore, it is important to provide accurate reliability predictions over time in order to determine accurately the repair, inspection and replacements strategies of products and systems.

Table 1
*The machine tool failure data**

Machine tool	Tool life (days)	Speed	Feed	Depth of cut
1	70	340	0.0063	0.021
2	29	570	0.0063	0.021
3	60	340	0.0141	0.021
4	28	570	0.01416	0.021
5	64	340	0.0063	0.021
6	32	570	0.0063	0.04
7	44	340	0.01416	0.04
8	24	570	0.01416	0.04
9	35	440	0.00905	0.029
10	31	440	0.00905	0.029
11	38	440	0.00905	0.029
12	35	440	0.00905	0.029
13	52	305	0.00905	0.029
14	23	635	0.00905	0.029
15	40	440	0.00472	0.029
16	28	440	0.01732	0.029
17	46	440	0.00905	0.0135
18	33	440	0.00905	0.0455
19	46	305	0.00905	0.029
20	27	635	0.00905	0.029
21	37	440	0.00472	0.029
22	34	440	0.01732	0.029
23	41	440	0.00905	0.0135
24	28	440	0.00905	0.0455

***Source**. Merrick and Soyer (2003)

1.2 PHM for machine tool failure time from non-Bayesian perspective

In this subsection, we briefly describe the problem and the model structure that we consider throughout the article. We follow Merrick and Soyer (2003) and let T_i be the life length of the machine tool i. Assuming that T_i is continuous, the failure rate function (failure rate is the frequency with which an engineered system or component fails, expressed for example in failures per hour, so it is important in reliability engineering) of the distribution of T_i is given by

$$\lambda_i(t) = \lim_{\Delta t \to 0} \frac{P(t \le T_i < t + \Delta t)}{\Delta t} = \frac{f_i(t)}{S_i(t)} \tag{2}$$

where $f_i(t)$ is the probability density function of T_i and

$$S_i(t) = P(T_i \ge t) = \exp\left(-\Lambda_i(t)\right) \tag{3}$$

is the reliability of machine tool i at time t, with $\Lambda_i(t) = \int_0^t \lambda_i(u)\, du$, the cumulative failure rate. The PHM for the data in Table 1 is specifying by

$$\lambda_i(t; Z_i) = \lambda_0(t) \exp\left\{\beta_1 \ln Z_{i,1} + \beta_2 \ln Z_{i,2} + \beta_3 \ln Z_{i,3}\right\} \tag{4}$$

The PH model is distribution-free requiring only the ratio of hazard rates between two stress levels to be constant with time, i.e., from (4)

$$\text{Hazard Ratio (HR)} = \frac{\lambda_i(t; Z_i)}{\lambda_0(t)} = \exp\left\{\beta_1 \ln Z_{i,1} + \beta_2 \ln Z_{i,2} + \beta_3 \ln Z_{i,3}\right\} \tag{5}$$

The dependent variable in PHM is time to failure (or survival time) of an observation. For the PHM, the coefficient of each variable is estimated using a partial likelihood method. Then the partial likelihood is maximized with respect to the parameters $\beta = (\beta_1, \beta_2, \beta_3)$. In most situations, we are interested in the parameter estimates than the shape of the hazard. The HR (5) , is the relative risk of failure at time t. From the interpretation of the model in the previous subsection it is obvious that β characterizes the 'effect' of Z. So β should be the focus of our inference while $\lambda_0(t)$ is "a nuisance parameter". Given a sample of survival data in Table 1, our inferential problems include:

1) Estimate β and derive its statistical properties;
2) Testing hypothesis H_0: $\beta = 0$;
3) Diagnostics.

Since the baseline hazard $\lambda_0(t)$ in (5) is left completely unspecified (infinite dimensional), ordinary likelihood methods can't be used to estimate β. In model (4), When T_i is a machine tool to right-censorship, we observe X_i = min (T_i, C_i) and $\delta_i = I(T_i \le C_i)$; where C_i is the censoring time and $I(E)$ indicates, by the values 1 versus 0, whether or not the event E occurs. Assume that T_i and C_i are independent conditional on Z_i. Let (X_i, δ_i , Z_i) ($i = 1,\ldots,n$) be independent observations. Andersen and Gill (1982) elegantly proved the asymptotic distribution of $\hat{\beta}$ by applying martingale theory in the counting process framework. In order to study the properties of the estimator $\hat{\beta}$, it is useful to formulate the problem by means of counting processes as follows:

Let us break the time axis (the follow-up time) into a grid of points. Assume the survival time is continuous. We hence can take the grid points dense enough so that at most one death can occur within any interval. Let $dN_i(t)$ denote the indicator for the i^{th} individual being observed to die in [t, t + t), such that

$$dN_i(t) = I(X_i \in [t, t + \Delta t),\ \delta_i = 1)$$

Let $Y_i(t)$ denote the indicator for whether or not the i^{th} individual is at risk at time t, such that $Y_i(t) = I(X_i \ge t)$. Let $dN(t) = \sum_{i=1}^{n} dN_i(t)$ denote the number of failures for the whole sample that occur in [t, t + t). Since, it is assumed that t is sufficiently small, so dN(t) is either 1 or 0 at any time t. The question now is, among $Y(t)$ = $\sum_{i=1}^{n} Y_i(t)$ individuals, what is the probability that the observed failure happened to the i^{th} machine tool (who is actually observed to fail at a point time t) rather than to the other machine tool. The estimation for β is the answer of that question. Parameter estimates in the Cox PHM are obtained by maximizing "the partial likelihood function", which derived by Cox (1975) derived the to be as

$$PL(\beta) = \prod_{i=1}^{n} \prod_{\{\text{all grid points t}\}} \left[\frac{\exp\{\beta' Z_i(X_i)\}}{\sum_{\ell=1}^{n} \exp\{\beta' Z_\ell(X_i)\}} \right]^{dN_i(t)} \tag{6}$$

$$PL(\beta) = \prod_{i=1}^{n} \left[\frac{\exp\{\beta' Z_i(X_i)\}}{\sum_{\ell=1}^{n} Y_i(X_i)\exp\{\beta' Z_\ell(X_i)\}} \right]^{\delta_i}$$

where $R_i(X_i)$ is the risk set, i.e., the set of indices corresponding to individuals at risk and uncensored at time X_i. The corresponding score function is:

$$U(\beta) = \frac{\partial}{\partial\beta}\log PL(\beta) = \sum_{i=1}^{n}\delta_i\left\{Z_i(X_i) - \frac{S^{(1)}(\beta,X_i)}{S^{(0)}(\beta,X_i)}\right\} \quad (7)$$

where $S^{(0)}(\beta,X_i) = \sum_{\ell\in R_i(X_i)}\exp\{\beta' Z_\ell(X_i)\}$ and $S^{(1)}(\beta,X_i) = \sum_{\ell\in R_i(X_i)} Z_\ell \exp\{\beta' Z_\ell(X_i)\}$.

The maximum partial likelihood estimator $\hat{\beta}$ is the solution to $U(\beta) = 0$. Tsiatis (1981) prove that $\sqrt{n}\left(\hat{\beta}-\beta_0\right)$ converges to multivariate normal distribution (with mean Zero and variance $I^{-1}(\beta)$) where

$$I = -E\left[\frac{\partial^2}{\partial\beta\partial\beta'}\log PL(\beta)\right]$$

Johnsen (1983) demonstrated that the partial likelihood (6) may be viewed as a “profile likelihood” in which the unknown baseline function $\lambda_0(t)$ is replaced in the total likelihood by a nonparametric maximum likelihood estimate. The PHM assumes that the effects of the $lnZ_i = (lnZ_{1,i}, lnZ_{2,i}, lnZ_{3,i})$ on hazard ratio are constant over time. In other words, it assumes that the hazard for one firm is a fixed proportion of the hazard for any other firm in the study. However, according to Atta and Sözer (2007), ignoring the non-proportional hazards in an analysis can lead us to incorrect results. For this reason, with applying the model (5) to the failure data in Table 1, one should first check the proportional hazards assumptions. The structure of this paper is as follows. Section 2, will present a Bayesian semi-parametric PHM to analyze machine tool failures in Table 1.

Section 3 shows how the method can be adapted reanalyzed using the classical or non-Bayesian (frequentist) approach and then turn to apply the Bayesian approach using WinBUGS software. The paper concludes with a discussion.

2 The proportional hazards model and Bayes structure

In the statistical literature many methods have been presented to deal with PHM with censored observations, both within the Bayesian and non-Bayesian frameworks, and such methods have been successfully applied to, e.g., reliability problems. Also, in reliability theory it is often emphasized that, through shortage of statistical data and possibilities for experiments, one often needs to rely heavily on judgments of engineers, or other experts, for which means Bayesian methods are attractive. It is therefore important that such judgments can be elicited easily to provide informative prior distributions that reflect the knowledge of the engineers well. In this Section, we will review the Cox PHM from Bayesian point of view.

2.1 Application of martingales in the PHM

The usage and success of Cox PHM are now a big achievement for the analysis of survival time data with covariates (Lawless, 1982). It turned out that the martingale concept had an important role to play in Cox model. Counting processes provided a natural framework in which to study the phenomenon and research activities in this area were already on the agenda, as exemplified above. In the original article for Cox (1972), he used a factor of the full likelihood, which he later justified and termed partial likelihood in Cox (1975), to estimate β. It soon became apparent that the Cox model could be immediately applied for the recurrent event intensity, and Johansen’s (1983) derivation of Cox’s partial likelihood as a profile likelihood also generalized quite easily. The analysis of counting process data, including survival data, is usually based on the modeling of the intensity.

Andersen and Gill (1982) extended (2.3) to the counting process framework and gave elegant martingale proofs for the asymptotic properties of the associated estimators in the models try to fit models for survival data. Others that have contributed to establishing asymptotic results for the model are Tsiatis (1981) and Næs (1982). If we have $n = 24$ machine tools under investigations, for the machine tool i, $i = 1,\ldots,24$; $I_i(t)$ is the intensity process for a counting process given covariate vector $Z_i = (Z_{i,1}, Z_{i,2}, Z_{i,3})$, and $Y_i(t)$ is the at risk indicator, i.e., the set of subjects still at risk at the time, T_i , of failure for subject i (i.e., alive and uncensored at time point just before time *t*), furthermore, we observe process $N_i(t)$, to count the number of failures which occurred in the interval [0, *t*]. That process is constant and equal to zero between failures and jumps one unit at each failure time. Hence, the rate of a new failure is then seen to be $I_i(t) = Y_i(t)\lambda(t \mid Z_i)$. The intensity $I_i(t)$ may be characterized as the probability that the event of interest occurs in the small time interval $[t, t+dt)$, given that it has not happened before. This gives approximately

$$dN_i(t) \approx \lambda(t \mid Z_i)dt = \lambda_0(t)\exp\left[\beta' Z_i\right] = I_i(t) \tag{8}$$

where $dN_i(t)$ is the increment of $N_i(t)$ over the small interval $[t, t+dt)$ (i.e., number of observed failures occurring in $[t, t+dt)$). Hence $I_i(t)$, is multiplicative intensity which can be modeled by

$$I_i(t) = Y_i(t)\lambda(t \mid Z_i) = Y_i(t)\lambda_0(t)\exp\left[\beta' Z_i\right], \tag{9}$$

where the intensity process is a product of an observed process and an unobserved function. Hence the intensity process for $N_i(t)$ under model (4) is:

$$I_i(t)dt = Y_i(t)\exp\left(\beta' Z_i\right)d\Lambda_0(t)$$

where $d\Lambda_0(t)$ represents the instantaneous probability that the subject at risk at time *t* has an event in the next time interval $[t, t+dt)$. As pointed out by Clayton (1991; 1994), it turned out that the martingale theory was of fundamental importance.

Suppose that the machine tools were followed to death or censored in a study. Thus we have observed data $D = \{N_i(t),\ Y_i(t),\ Z_i;\ i = 1, 2, ..., n\}$ and we have unknown parameters β, $\Lambda_0(t)$. Under non-informative censoring, the likelihood of the data factorizes, with one term depending only on the censoring process and the second term:

$$L_i\left(D \mid \beta, \Lambda_i(t)\right) = \exp\left(-\int_{t\geq 0} I_i(t)dt\right)\prod_{t\geq 0}\left[I_i(t)\right]^{dN_i(t)};\ i = 1, 2, ..., n. \tag{10}$$

Hence, according to the probability model under consideration, the variables *D* have their joint distribution given by

$$L\left(D \mid \beta, \Lambda(t)\right) = \prod_{i=1}^{n} L_i\left(D \mid \beta, \Lambda_i(t)\right) \tag{11}$$

Following the counting process notation introduced above, for machine tool *i* ($i = 1, 2, ..., n$) represent the process counting the failures occurring up to time *t*, while $dN_i(t)$ is a small increment of $N_i(t)$ over the interval [*t*, *t* + *dt*). $N_i(t)$ and $dN_i(t)$ equal 1 if the event occurs in [0, *t*) and [*t*, *t* + *dt*), respectively, and 0 otherwise. Under non-informative censoring, the observed counting process likelihood is proportional to (10) is following a Poisson form. Though $dN_i(t)$ is at most one for all i, t; the infinitesimal counting process increments, $dN_i(t)$, contribute to the likelihood just as with independent Poisson random variables with means $I_i(t)dt$ over the interval $[t, t+dt)$. Defining the model in this framework allows the intensity to be regarded as constant in that interval (Clayton, 1991). For efficient posterior computations, we implement a data augmentation approach based on the consideration that $dN_i(t)$ are independent Poisson random variables under the likelihood expression (11), i.e.

$$dN_i(t) \sim \text{Poisson}\left(I_i(t)dt\right) \tag{12}$$

After D is observed, our interest turns to the posterior distribution $P\left(\beta, \Lambda_0(t) \mid D\right)$. Baye's Theorem tells us that $P\left(\beta, \Lambda_0(t) \mid D\right) \propto P\left(D, \beta, \Lambda_0(t)\right)$. Thus, whole probability model can now be expressed as the joint posterior distribution of the model parameters according to the Bayesian approach, i.e.

$$P\left(\beta, \Lambda_0(t) \mid D\right) \propto L\left(D \mid \beta, \Lambda_0(t)\right)P(\beta)P\left(\Lambda_0(t)\right) \tag{13}$$

The focus is on inference for β. Another quantity of interest is the baseline hazard function, $\Lambda_0(t)$ which is best viewed as a process over time. As (13) has a complicated form, it is most conveniently summarized using simulation. Thus might be accomplishing using Gibbs sampler. Implementing a Gibbs sampler coded from scratch would require us to identify and then construct an effective simulation method for each of the two related full conditional posterior distributions, namely $P(\beta \mid D, \Lambda_0(t))$, $P(\Lambda_0(t) \mid D, \beta)$. The Cox's partial likelihood (6) can be viewed as a limiting case of the marginal posterior of β in the Cox model with a gamma process prior on the cumulative baseline hazard (See Chapter 3.2.3, Ibrahim et al., 2001).

2.2 Prior specification and inference

The primary motivation of using Bayesian statistics in reliability analysis is the ability to incorporate prior knowledge with limited testing results in a formal procedure. This idea is particularly suitable for high reliable systems, which cannot afford enough samples to meet the confidence requirement in reliability demonstration test. The problem of appropriate choice of prior distribution is the central one confronting the reliability users of Bayesian methods. Specifying the model (13) described above will allow us to make use of a user-friendly package for Gibbs sampling such as WinBUGS (see, e.g., Lunn et al., 2000; Spiegelhalter et al., 2003). In WinBUGS, we assigned a priori distributions to the coordinates of β using independent normal distributions with mean Zero and a large variance such as 10,000. This reflects our lack of knowledge the risk factors would increase or decrease failure time. Approaches for modeling the prior belief for the baseline cumulative failure rate $\Lambda_0(t)$ will be adopted to be the gamma process in our analysis, which was suggested by Kalbfleisch (1978); hence we let

$$\Lambda_0(t) \sim GP\left(c_0\Lambda^*(t), c_0\right);$$

where $\Lambda^*(t)$ is the mean of the process such that $\Lambda^*(t)$ is increasing function with $\Lambda^*(t) = 0$ and c_0 is a specification of weight or confidence attached to that guess (a weight parameter of the mean (Ibrahim et al., 2001)). Kalbfleisch (1978) showed that if $c_0 \approx 0$; then the likelihood (13) is approximately proportional to the partial likelihood (6), whereas $c_0 \to \infty$, the limit of the likelihood when the Gamma process is replaced by $\Lambda^*(t)$. In general, $\Lambda^*(t)$ is taken to be a known parametric function, such as exponential or Weibull distribution. For example, if specified vector of hyper-parameters. The values of the hyper-parameters should be carefully selected to avoid a convergence problem or floating problem in MCMC sampling.

3 Comparing Bayesian and non-Bayesian frameworks results for EIS

To compare between Bayesian and non-Bayesian (frequentist) results in fitting hazard ratio (5), our basic strategy in this Section will be devoted to find the parameters of interest in Cox PHM from non-Bayesian approach and then turn to obtain the posterior distribution of the regression parameters and the survival function using a combination of recent Monte Carlo methods. We will use the dedicated Bayesian software WinBUGS that implement MCMC methodology. BUGS is an acronym for '**B**ayesian inference **U**sing **G**ibbs **S**ampling'. Gibbs sampling is a specific MCMC method. The Gibbs sampler yields a Markov chain whose stationary distribution is the posterior distribution. An illustrative analysis within the data analyzed by Merrick and Soyer (2003), listed in Table 1 as reported at the end of Section 1, consist of 'the time to occurrence failure' for each machine in 24 ones. The dimension of covariate vector is 3 with $Z_i = (Z_{i,1}, Z_{i,2}, Z_{i,3})$; $Z_{i,1}$ is the cutting speed, $Z_{i,2}$ is the feed rate and $Z_{i,3}$ is the depth of cut. Moreover, we show how the output of the MCMC algorithm can be used to obtain draws from the posterior of parameter distributions.

3.1 Application to an engineering insurance system using frequentist statistics

First, the Cox PH model is fitted to the data in Table 1. The Global Null Hypothesis: $\beta = 0$ (the null hypothesis is that none of the risk factors has a statistically significant influence on the hazard rate (4)). Three statistical testing procedures, Likelihood ratio test, Score, and Wald, were applied. The results showed that there is a less than 0.01% (which is highly significant) chance that one will make a faulty rejection of the null hypothesis. So, the data strongly suggests that the hazard rate is dependent on the selected risk factors (cutting speed, feed rate and depth of cut). Thus we conclude that the three explanatory variables are needed to adjust the lifetime test. The results of the Cox PH model are summarized in Table 2, giving the estimators of hazard ratios for each covariate and their confidence intervals and its *p*-value from the likelihood ratio test.

Table 2

Cox's proportional hazards analysis for the machine tool failure data

	β	SE	*p*-value	Hazard ratio	95% conf. interval	
lnZ1	12.131	2.542	1.3×10^{-6}	185519.329	1272.064	2.706×10^{7}
lnZ2	1.975	.746	0.005625	7.209	1.671	31.097
lnZ3	3.001	.899	0.000935	20.101	3.449	117.146

A point estimate of the effect of β is provided in the Hazard ratio (the slope of the lifetime curve that is a measure of how rapidly machine tools are failing) is given column 5 in Table 2. The cutting speed, the feed rate and the depth rate are the three covariates that shows a statistically significant impact on the reliability (survival) at the level of $\alpha = 0.05$. Thus, after the final model of significant explanatory variables was created, it was necessary to validate the proportional hazards assumption. The proportional hazards model hypotheses are tested for each covariate based on scaled Schonfeld residuals (see, Therneau and Grambsch, 2000). In Figure 1 the plot of scale residuals are given against order time along with spline smooth together 95% confidence intervals. The broken lines represent ± 2-standard-error envelopes around the fit. The plot does not show a strong trend along the original time variable and the systematic departures from a horizontal line are indicative of non-proportional hazards. Furthermore, the proportional hazards assumption was assessed by the Score test, in Table 3; the first column is the correlation of the scaled Schoenfeld residuals with the time variable. The second column is the test statistic defined previously. The global test is to test simultaneously all the slopes are zero. All the *p*-values are fairly large, indicating that the slopes are zero. In conclusions, through the model diagnostics, we find that the model (4) considered fits the data in Table 1 very well.

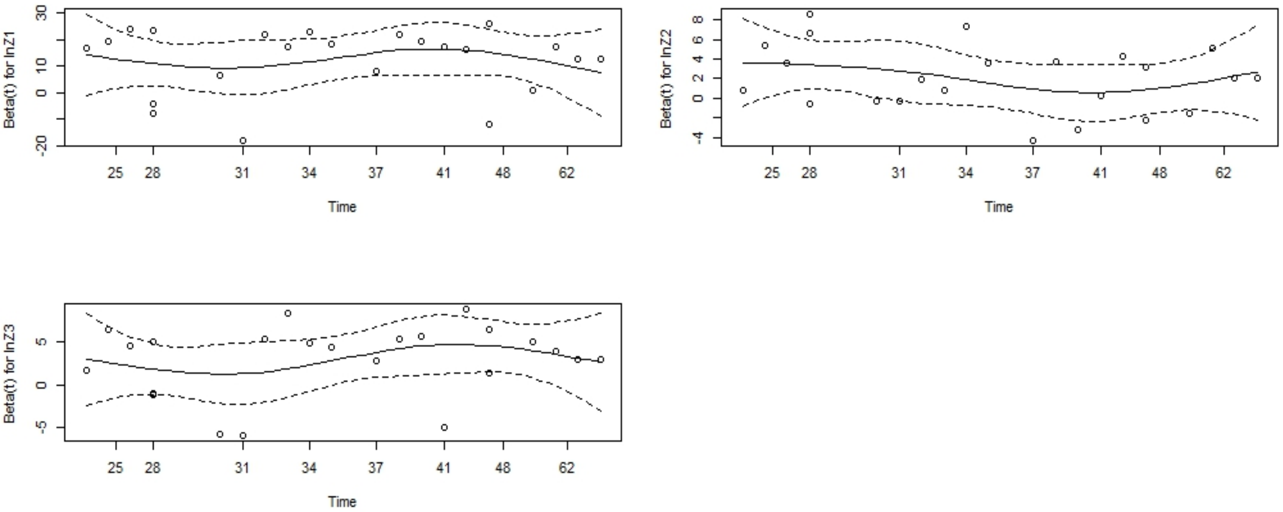
Table 3

Score test of proportional hazards assumptions

Variables	Rho	Chi-Square	*p*-value
lnZ 1	0.0271	0.016	0.899
lnZ 2	−0.2395	1.049	0.306
lnZ 3	0.1865	0.694	0.405
Global test	.	2.376	0.498

Fig.1

A graphical assessment of the proportional hazards assumption: Plots of scaled Shoenfeld residual against ordered time with a spline smooth for the three covariates.



Influential observations can be detected by investigating the so-called "DfBeta", diagnostics, of estimated changes in the regression coefficient upon deleting each observation in turn; that can computed for each subject and covariate. This statistics measures the change in regression coefficient of the corresponding covariate, if a machine is left out from model estimation. If this value is high (positive or negative), the machine has more influence on the results than other machines. Investigation of the characteristics of such machine could lead to the detection of new risk factors. Plotting the DfBeta values for each covariate against machine tool as shown in Figure 3, we learn that, there are not any machine tool has noticeably identified as influential point. The martingale residual plot to check functional form of the covariate is plotted in Figure 4, for the three diagnostic factors. Figure 4 indicates a linear form seems appropriate for the three covariates, hence a dichotomized transformation of any covariates not needed (i.e., there is no evidence of nonlinearity here)

Fig. 3

Index plots of DfBeta for the Cox regression of time to failure on lnZ1, lnZ2 and lnZ3

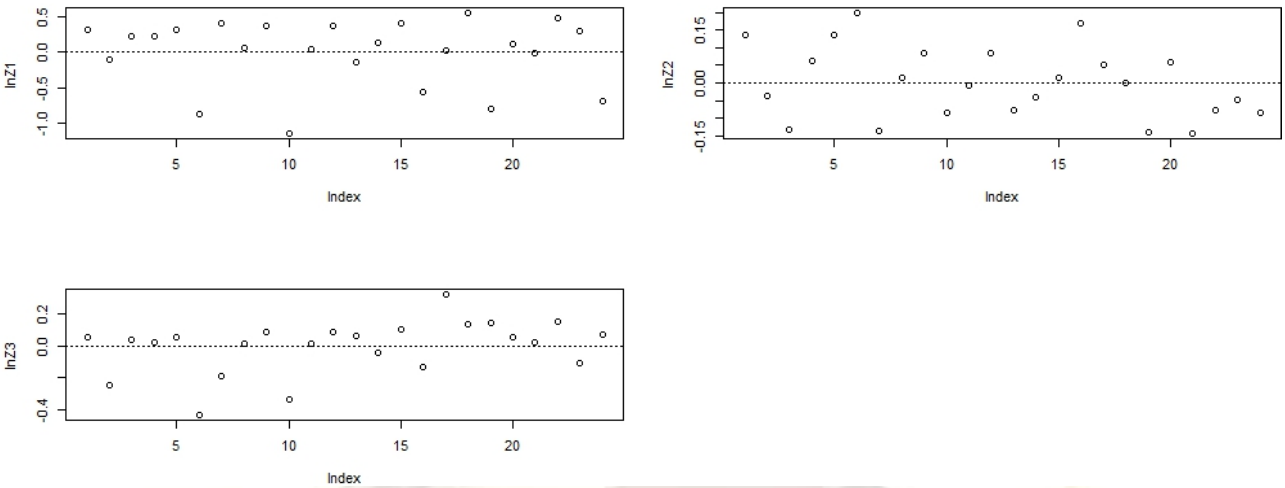
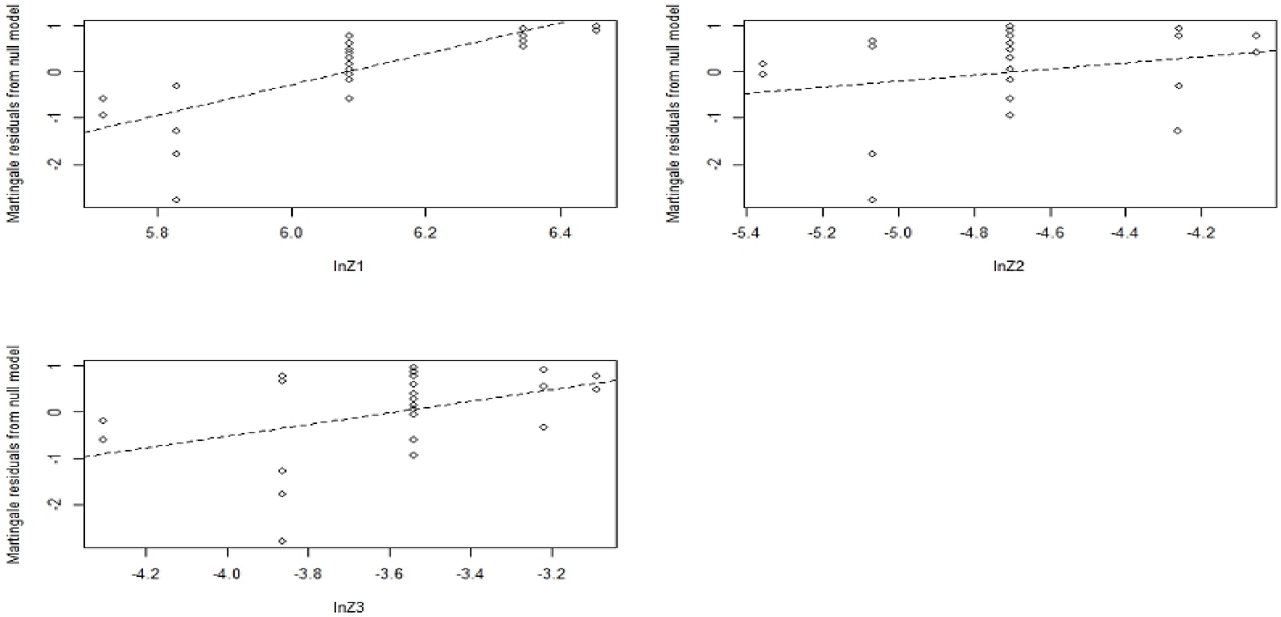


Fig. 4

Martingale residuals to check the functional form of the continuous variables lnZ1, lnZ2 and lnZ3



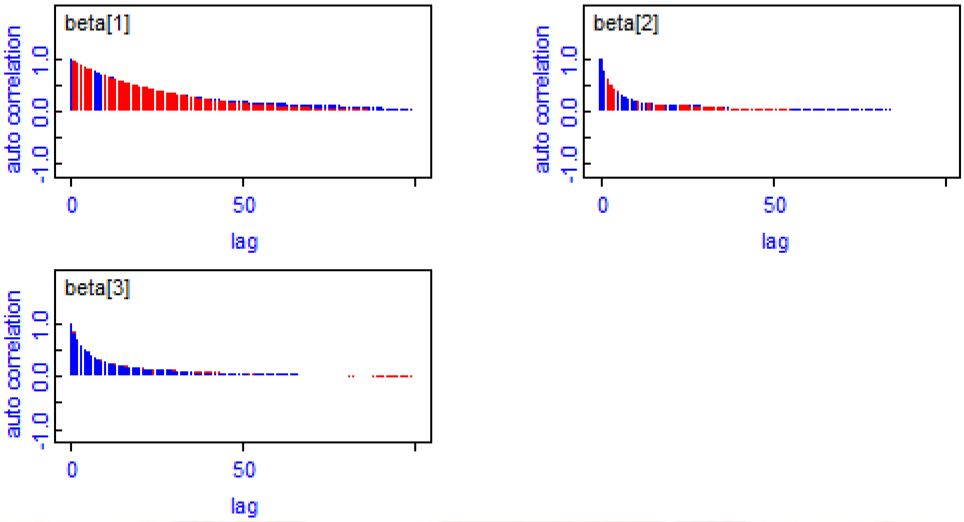
3.2 Application to an engineering insurance system using Bayesian statistics

As discussed above, the Cox PHM evaluates risk factors to determine the magnitude and significance of their effects on failure time. Here, we reanalyze the data listed in Table 1 using WinBUGS software. Given the model assumptions, this program performs the Gibbs sampler by simulating from the full conditional distributions. The Bayesian estimators were obtained through the implementation of the Gibbs sampling scheme described in Section 2. We implemented 40,000 iterations of the algorithm and described the first 1,000 iterations as a burn-in. Spiegelhalter et al., (2003), the BUGS team, use the idea of parallel multiple chains to check the convergence of the Gibbs sampler, we used 2 chains, as shown in Figure 5.

ISSN: 2248-9622 www.ijera.com
Vol. 2, Issue 2,Mar-Apr 2012, pp.374-385

Fig. 5

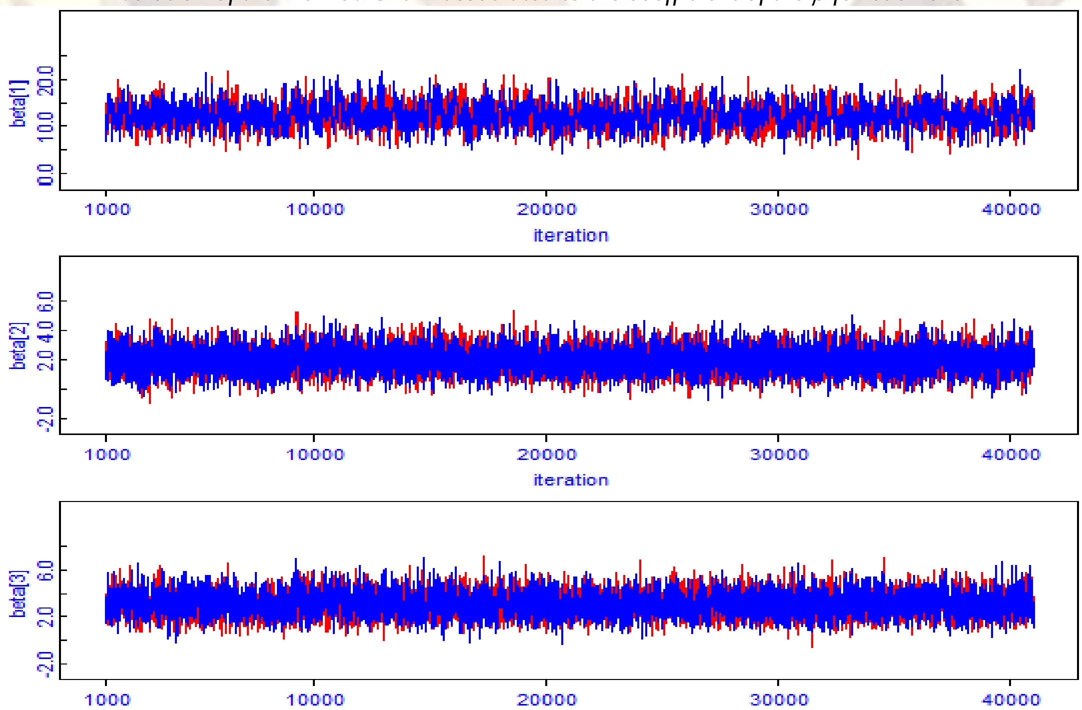
ACF for the iterations for each chain



The fully quantitative monitoring of parallel multiple chains was first proposed by Gelman and Rubin (1992a, b). The chains should start from over-dispersed initial values to ensure good converge of parameter space. To generate the Gibbs posterior samples in the previous section, we choose to use two parallel chains. Hence, once convergence has been achieved, 40,000 observations are taken from each chain after the burn-in period to reach our goal of 80,000 observations. The BUGS software offers also a graph of the autocorrelation function (ACF) of the iterations to the 50-lag for each chain independently (Figure 5). The autocorrelation plot in Figure 5 illustrates such dependence between successive observations, which appears to die out before lag 40. This indicates fairly rapid mixing and thus good convergence of the parameter space with a reasonably small number of iterations. As a result of thumb if the autocorrelations are needed to get rid of the dependence structure, but from (Figure 6), we can be reasonably confident that convergence of β has been achieved (the two chains appear to be overlapping one another) and thus the convergence looks reasonable. Finally, all the hyper-parameters of the model showed an acceptable convergence to the stationary distribution both with graphical tests, like the one in Figure 6, and with statistical tests like the Geweke one and the Raftery and Lewis test.

Fig. 6

Evolution of the Markov Chain associated to the coefficient of the β for each chain



In Table 4, 2.5% and 95.5% correspond to the respective posterior percentiles of β is obtained. Therefore, the 95% credible interval for β_1 is thus (7.89, 18.08), and the mass for the posterior distribution of β_1, β_2, β_3 are to the left of zero, indicating $Z_{i,1}, Z_{i,2}, Z_{i,3}$ have pronounce effect. This can be further illustrated in a plot of the marginal posterior density of β as shown below in Figure 7. When comparing Table 4 with Table 2, we draw the same conclusions about the parameter of interest.

Table 4.

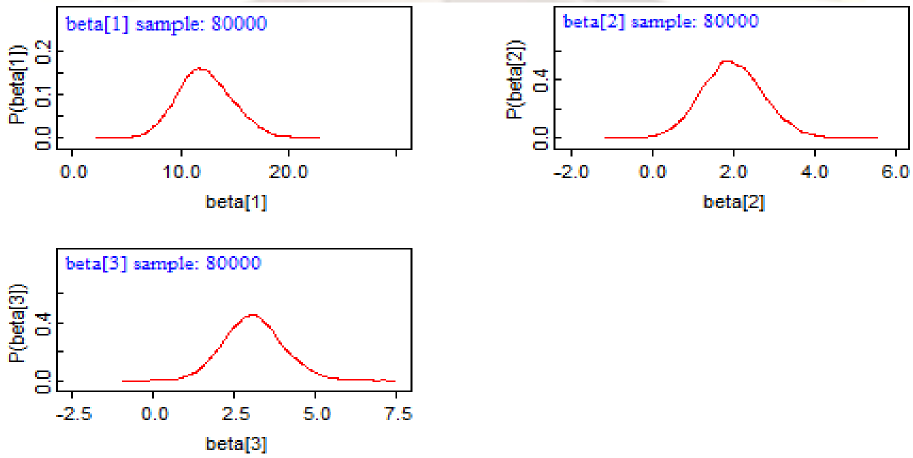
Posterior summary: ETS lifetime parameters

Parameter	Average	SE	25^{th} percentile	79.5^{th} percentile	Median
β_1	12.66	2.606	7.89	18.08	12.6
β_2	2.062	0.7637	0.6253	3.599	2.036
β_3	3.145	0.912	1.4	4.965	3.135

Fig. 7

Estimated marginal density for β

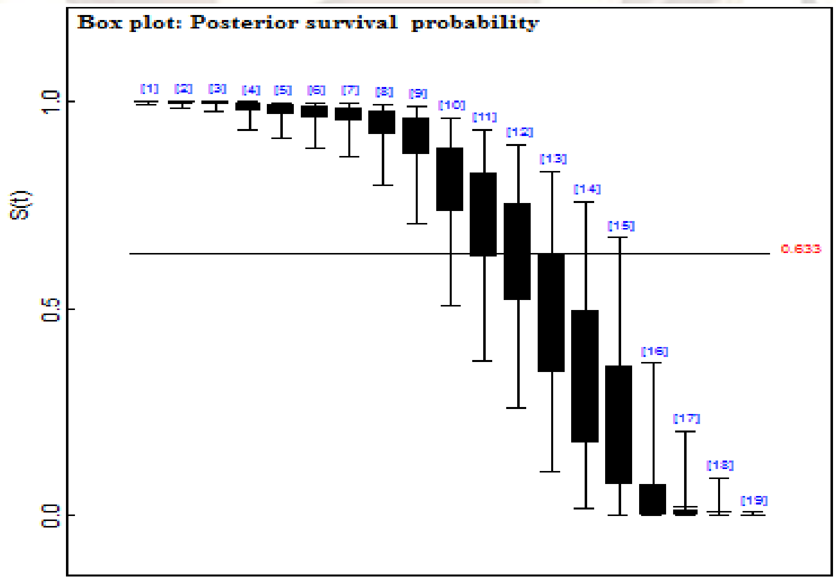
The posterior predictive distribution of the reliability of a given machine tool at defined mission times are shown as Box



plot in Figure 8 the semi-parametric model (4). The new machine tool is predicted to operate with a cutting velocity of 440 fpm, a feed rate of .00472 ipr, and a .029-inch depth of cut.

Fig. 8

Boxplots of the Posterior Distribution of the Survival Probability at Fixed Lifetimes Under the Parametric Model.



4. Conclusion and recommendation

Hazard ratios are a specific type of relative risk that is calculated using a statistical technique known as lifetime data analysis. The proportional hazard model in lifetime analysis was proposed by Cox (1972) is an ideal tool for formulating the relationship between event risks and their associated factors. When the underlying data is properly formatted, the estimated parameters, or coefficients, of the model provide intuitive measurements of risk variations for a given factor. This evaluation will very useful for any insurance company to determine its premiums in the insurance process based on these methods. We have used a semi-parametric model for survival populations based on the Cox proportional hazard model from both non-Bayesian and Bayesian perspectives. Such analysis was inconceivable only a few years ago. But now, with the great increase in computational power and memory capabilities of new computers, the analysis of these kinds of models is not only possible but advisable because they can better explain the relationship between the dependent variable and risk factors.

Acknowledgements

This research was supported by (Project #321212 through Deanery of Academic Research Fund), awarded to (Faculty of Sciences) in the Population Research Unit at The University of Al-Imam Muhammad Ibn Saud Islamic by Deanery of Academic Research.

References

[1] Andersen, P. K. and Gill, R. D., (1982). Cox’s regression model for counting processes: A large sample study. Annals of Statistics, 10, 1100–1120.
[2] Ata, N. and Sözer, M. T., (2007). Cox regression models with nonproportional hazards applied to lung cancer survival data. Hacettepe Journal of Mathematics and Statistics. 36, 2, 157-167. Available http://www.mat.hacettepe.edu.tr/hjms/english/issues/vol36_2.html
[3] Chen, K.C.W. and Lee, C.W.J., (1993). Financial ratios and corporate endurance: a case of the oil and gas industry. Contemporary Accounting Research, 9, 667–694.
[4] Clayton, D. A., (1991). Monte Carlo for Bayesian inference in frailty models. Biometrics, 47, 467– 485.
[5] Clayton, D. A., (1994). Bayesian analysis of frailty models. Technical Report, Medical Research Council Biostatistics Unit, Cambridge.
[6] Cox, D. R., (1975). Partial Likelihood. Biometrika, 62, 269–276.
[7] Cox, D.R., (1972). Regression models and life tables (with discussion). Journal of the Royal Statistical Society, Series B, 34: 187–220.
[8] Cox, D.R., and Oakes, D., (1984). Analysis of Survival Data, London: Chapman and Hall.
[9] Gelman, A. and Rubin, D. (1992a). Inference from iterative simulation using multiple sequences. Statistical Science, 7, 457-511.
[10] Gelman, A. and Rubin, D. (1992b). A single from the Gibbs sampler provides a false sense of security. Bayesian Statistics, 4, eds. J.M. Bernardo, J.O. Berger, A.P. Dawid and A.F.M. Smith, New York: Oxford University Press, 625-631.
[11] Ibrahim, J. G., Chen, M.-H., and Sinha, D. (2001). Bayesian Survival Analysis. Springer.
[12] Johansen, S., (1983). An extension of Cox’s regression model. International Statistical Review, 51, 165 – 174.
[13] Kalbfleisch, J. D., (1978). Nonparametric Bayesian analysis of survival time data. Journal of the Royal Statistical Society, Series B, 40, 214 – 221.
[14] Keiding, N., (1998). The Cox regression model for claims data in non-life insurance. Astin Bulletin, 28, 1, 95–118. Available http://www.casact.org/library/astin/vol28no1/
[15] Kiefer, N.M., (1988). Economic duration data and hazard functions. Journal of Economic Literature, 646 –679.
[16] Kim, Y., Anderson, D.R., Amburgey, T.L. and Hickman, J. C., (1995). The use of event history analysis to examine insurer insolvencies. Journal of Risk and Insurance, 62, 94–110.
[17] Lane, W.R., Looney, S.W., and Wansley, J.W., (1986). An Application of the Cox Proportional Hazards Model to Bank Failure. Journal of Banking and Finance, 10, 511–531.
[18] Lawless, J. F., (1982). Statistical Models and Methods for Lifetime Data. John Wiley and Sons, New York.
[19] Lunn, D., Thomas, A., Best, N. G., and Spiegelhalter, D., (2000). WinBUGS — A Bayesian modelling framework: Concepts, structure, and extensibility. Statistics and Computing 10, 325–337.
[20] Merrick, J. R. and Soyer, R., (2003). A Bayesian semiparametric analysis of the reliability and maintenance of machine tools. Technometrics, 45, 1–12. Available http://home.gwu.edu/~soyer/
[21] Mostafa, A. A. and Ben Ghorbal, A., (2011). Using WinBUGS to Cox model with changing from the Baseline Hazard Function Applied Mathematical Sciences, 45, 2217 – 2240. Available http://www.m-hikari.com/ams/ams-2011/ams-45-48-2011/
[22] Næs, T., (1982). The asymptotic distribution of the estimator for the regression parameter in Cox’s regression model. Scandinavian Journal of Statistics, 9, 107–115.

[23] Ng, S.H., Cram, F., and Jenkins, L., (1991). A Proportional hazards regression analysis of employee turnover among nurses in New Zealand. Human relations, 44, 1313–1329.

[24] Spiegelhalter, D. J., Thomas, A., Best, N. G., and Lunn, D., (2003). WinBUGS 1.4 User manual.

[25] Therneau, T. M. and P. M. Grambsch., (2000). Modeling Survival Data: Extending the Cox Model. New York: Springer.

[26] Tsiatis, A. A., (1981). A large sample study of Cox's regression model. Annals of Statistics, 9, 93 – 108.

[27] Whalen, G., (1991). A Proportional hazards Model of bank failure: an examination of its usefulness as an early warning tool. Federal Reserve Bank of Cleveland, 21–31.

[28] Xiong, C. Zhu, K. and Yu, K., (2006). Statistical Survival Analysis with Applications. Springer.