Journal of Statistical
Theory & Methods
1996, Vol. 7, pp 227~234

Discount Survival Models

Joo Y. Shim and Joong K. Sohn

Abstract The discount survival model is proposed for the application of the Cox model on the analysis of survival data with time-varying effects of covariates. Algorithms for the recursive estimation of the parameter vector and the retrospective estimation of the survival function are suggested. Also the algorithm of forecasting of the survival function of individuals of specific covariates in the next time interval based on the information gathered until the end of a certain time interval is suggested.

Keywords : Discount Model, Cox Model, Survival Data, Covariate, Recursive Estimation, Survival Function.

1. Introduction

Cox(1972) proposed the regression model for the analysis of survival data by introducing a parameter vector modeling effects of covariates in the model under the assumption that covariates have fixed effects on the survival pattern. West *et al.*(1985) developed the dynamic generalized linear model, where the distribution of a time-varying parameter is partially specified by its mean and variance-covariance matrix. Ameen and Harrison(1985) developed the normal discount Bayesian model to overcome some practical disadvantages of the dynamic linear model. They introduced the use of a discount factor d $(0<d<1)$ so that the variance of a parameter of the current time is equal to the variance of a parameter of the previous time divided by a discount factor, which implies an increase in variance of $100(1-d)/d$ percent. Gamerman(1991) proposed the dynamic Bayesian model for the analysis of survival data with a time-varying parameter vector. It is assumed that the time-variation of the parameter vector is determined through the evolution of the parameter vector between time intervals.

The discount survival model is proposed by incorporating the Cox model into the dynamic generalized linear model and the discount Bayesian model. This model provides a quick response to sudden changes of the time-varying parameter vector, which leads a faithful representation of survival data via survival functions.

The discount survival model is described in Section 2. The procedure of recursive estimations of the parameter vector under the multiprocess dynamic model is provided in Section 3. Also the procedure of retrospective estimations of the survival function

Joo Y. Shim and Joong K. Sohn are a lecturer and a professor, respectively, in the Department of Statistics, Kyungpook National University 702-701, Taegu, Korea.

and the procedure of one-step ahead forecasting of survival functions in the next time interval are provided in Section 4 and Section 5, respectively. In Section 6 the performance of the proposed model is illustrated via the real data.

2. Model Description

Here we assume that the survival time has a piecewise exponential distribution which has a constant hazard rate in each time interval,

$$\lambda(\tau) = \lambda_i \quad \text{for} \quad t \in I_i = (\tau_{i-1}, \tau_i]$$

where τ_0 is usually set to 0 and $I_s = (\tau_{s-1}, \infty)$. The survival function and the probability function for the current time interval I_i given survival up to the end of the previous time interval can be easily calculated due to the lack of memory property of exponential distributions. Thus we obtain the likelihood for λ_i under the assumption that the random censoring time has no relation with the survival time,

$$l(\lambda_i|t) = \lambda_i^{\delta} \exp(-\lambda_i(t - \tau_{i-1})) \quad \text{for} \quad t \in I_i$$

where δ is the indicator function of a death of the individual in the time interval I_i. We denote the hazard rate corresponding to the *j-th* individual alive at the beginning of time interval I_i by $\lambda_{i(j)}$, $j = 1, \cdots, n_i$, where n_i is the number of individuals used to be alive at the beginning of the time interval I_i. Let T_{ij} be a survival time of the *j-th* individual used to be alive at the beginning of the time interval I_i. Also let t_{ij} be an observed failure time corresponding to T_{ij}, which can be interpreted as the minimum of survival time of the *j-th* individual and corresponding censoring time. We define the information about the *j-th* individual as the information of the survival time of the *j-th* individual alive at the beginning of time interval I_i, which is obtained at time t_{ij}, consisting of following 3 events;

i) the *j-th* individual dies at t_{ij}
ii) the *j-th* individual is censored at t_{ij}
iii) the *j-th* individual is alive at $t_{ij} = \tau_i$ so that $T_{ij} > \tau_i$. In first two cases provide no further information of the survival time of the *j-th* individual.

Let D_i be a set of informations from all observations of time intervals, $I_1, \cdots, I_i$, and let $D_{i-1(j)}$ be a set containing D_{i-1} and informations from first j observations of time interval I_i so that $D_{i-1(0)} = D_{i-1}$ and $D_{i-1(n\_i)} = D_i$. Then the discount survival model in the time interval I_i is defined by

i) observation equation ;

$$T_{ij} \sim exp(\lambda_{i(j)}, I_i) \quad \text{for} \quad j = 1, \cdots, n_i, i = 1, \cdots, s,$$

ii) guide relationship;

$$\lambda_{i(j)} = \exp(Z_{ij}' \beta_i) \quad for \ j = 1,\cdots,n_i, i = 1,\cdots,s,$$

iii) evolution equation;

$$E(\beta_i | D_{i-1}) = E(\beta_{i-1} | D_{i-1})$$
$$V(\beta_i | D_{i-1}) = B_i V(\beta_{i-1} | D_{i-1}) B_i, \ \text{for } i = 1,\cdots,s.$$

where B_i is a diagonal matrix of discount factors in the time interval I_i.

3. Recursive Estimation of a Parameter Vector

The process is started with the initial distribution of β_0 specified by mean vector $\hat{\beta}_0$ and variance-covariance matrix V_0, where $\hat{\beta}_0$ and V_0 are given prior to the time interval I_1, which do not affect distributional behaviors of the parameter vector in future time intervals for a certain extent of time elapsing.

At the beginning of each time interval I_i, the posterior distribution of β_{i-1} obtained in time interval I_{i-1} leads the prior distribution of β_i

$$(\beta_i | D_{i-1}) \sim [a_i, R_i]$$

where

$$a_i = \hat{\beta}_{i-1} \text{ and } R_i = B_i V_{i-1} B_i.$$

With informations from first *(j-1)* observations, the joint prior distribution of β_i and $\log \lambda_{i(j)}$ is obtained by the guide relationship,

$$\left(\begin{matrix} \beta_i \\ \log \lambda_{i(j)} \end{matrix} \middle| D_{i-1(j-1)} \right) \sim \left[\begin{pmatrix} a_{ij} \\ f_{ij} \end{pmatrix}, \begin{pmatrix} R_{ij} & S_{ij} \\ S_{ij}' & q_{ij} \end{pmatrix} \right] \tag{3.1}$$

where

$$f_{ij} = Z_{ij}' a_{ij}, \ S_{ij} = R_{ij} Z_{ij} \ \text{and} \ q_{ij} = Z_{ij}' S_{ij},$$

with $a_{i1} = a_i$ and $R_{i1} = R_i$. Here the prior distribution of $\lambda_{i(j)}$ is assumed to be a conjugate gamma distribution (b_{ij}, r_{ij}), where b_{ij} and r_{ij} are estimated in terms of the mean and the variance of the distribution of $\log \lambda_{i(j)}$, such as, respectively, q_{ij}^{-1} and $q_{ij}^{-1} \exp(-f_{ij})$.

With information from the *j-th* observation the posterior distribution of $\lambda_{i(j)}$ is obtained as

$$(\lambda_{i(j)} | D_{i-1(j)}) \sim Ga(b_{ij} + \delta_{ij} \ , \ r_{ij} + t_{ij} - t_{i-1}). \tag{3.2}$$

Using (3.2) and applying the linear Bayes estimation on (3.1) the updated distribution of β_i given $D_{i-1(j)}$ is estimated as

$$(\beta_i | D_{i-1(j)}) \sim [\hat{\beta}_{ij}, V_{ij}]$$

where

$$\hat{\beta}_{ij} = a_{ij} + S_{ij} q_{ij}^{-1} \log\left(\frac{1 + q_{ij}\delta_{ij}}{1 + q_{ij}(t_{ij} - \tau_{i-1})\exp(f_{ij})} \right),$$

$$V_{ij} = R_{ij} + S_{ij} S_{ij}' \left(\frac{\delta_{ij}}{1 + q_{ij}} \right).$$

Since there is no parametric evolution in each time interval, the joint prior distribution of β_i and $\log \lambda_{i(j+1)}$ is given as

$$\left(\begin{matrix} \beta_i \\ \log \lambda_{i(j+1)} \end{matrix} \middle| D_{i-1(j)} \right) \sim \left[\begin{pmatrix} a_{i,j+1} \\ f_{i,j+1} \end{pmatrix}, \begin{pmatrix} R_{i,j+1} & S_{i,j+1} \\ S_{i,j+1}' & q_{i,j+1} \end{pmatrix} \right]$$

where

$$a_{i,j+1} = \hat{\beta}_{ij}, \quad f_{i,j+1} = Z'_{i,j+1} a_{i,j+1}$$
$$R_{i,j+1} = V_{ij}, \quad S_{i,j+1} = R_{i,j+1} Z_{i,j+1}$$
$$q_{i,j+1} = Z_{i,j+1}' S_{i,j+1}.$$

Thus, when all individuals in time interval I_i are observed, the posterior distribution of β_i is estimated as

$$(\beta_i | D_i) \sim [\hat{\beta}_i, V_i]$$

where

$$D_i = D_{i-1(n_i)}, \quad \hat{\beta}_i = \hat{\beta}_{in_i}, \quad \text{and} \quad V_i = V_{in_i}.$$

4. Estimation of the Survival Function

In this section, under the assumption of the conjugate gamma distribution of the corresponding hazard rate, the estimated survival function of individuals of a covariate vector Z_h is obtained with informations gathered until the end of time interval I_N.

Using smoothed distribution of β_i given D_N, the distribution of the corresponding hazard rate given D_N is obtained by the guide relationship, which leads to the estimation of the survival function. Through the evolution equation, the joint distribution of β_i and β_{i+1} given D_i is obtained as

$$\left(\begin{matrix} \beta_i \\ \beta_{i+1} \end{matrix} \middle| D_i \right) \sim \left[\begin{pmatrix} \hat{\beta}_i \\ a_{i+1} \end{pmatrix}, \begin{pmatrix} V_i & V_i \\ V_i & R_{i+1} \end{pmatrix} \right]. \tag{4.1}$$

Using the distribution of β_{i+1} given D_i and applying the linear Bayes estimation on

(4.1), the smoothed distribution of β_i given D_N is estimated as

$$(\beta_i|D_N)\sim[\hat{\beta}_{i:N}, V_{i:N}], \tag{4.2}$$

where

$$\beta_{i:N} = \hat{\beta}_i + V_i G'_{i+1}R_{i+1}(\hat{\beta}_{i+1:N} - a_{i+1}),$$
$$V_{i:N} = V_i - V_i R_{i+1}(R_{i+1} - V_{i:N}R_{i+1}V_i).$$

Applying the guide relationship on (4.2) estimates of the defining parameters of the assumed distribution of $\lambda_{i(h)}$, $Ga(b_i, r_i)$ are obtained as, respectively, $(Z'_h V_{i:N}Z_h)^{-1}$ and $\exp(-Z'_h \hat{\beta}_{i:N})(Z'_h V_{i:N}Z_h)^{-1}$.

By integration

$$P(T_h > t|T_h > \tau_{i-1}, D_N) = \left(1 + \frac{t - \tau_{i-1}}{r_i}\right)^{-b_i}.$$

Thus, at the end of time interval I_N, by Bayes rule, the survival function for individuals of the covariate vector Z_h is obtained as

$$P(T_h > t|D_N) = \left(1 + \frac{t - \tau_{i-1}}{r_i}\right)^{-b_i} \prod_{k=1}^{i-1}\left(1 + \frac{\tau_k - \tau_{k-1}}{r_k}\right)^{-b_k}.$$

5. Forecasting of the Survival Function

In this section, the one-step ahead forecasted survival function of individuals of the specific covariate vector z_h used to be alive at the end of time nterval I_i is estimated based on informations gathered until the end of interval I_i.

At the end of a specific time interval I_i, through which individuals of the covariate vector Z_h has been observed alive, the posterior distribution of β_i which is equivalent to the distribution of β_i given $T_h > \tau_i$ and D_i is estimated as

$$(\beta_i|T_h > \tau_i, D_i)\sim\lfloor\hat{\beta}_i, V_i\rfloor.$$

By the evolution equation and the guide relationship, the joint prior distribution of β_{i+1} and $\log\lambda_{i+1(h)}$ is obtained as

$$\begin{pmatrix} \beta_{i+1} \\ \log\lambda_{i+1(h)} \end{pmatrix}|T_h > \tau_i, D_i\Bigg) \sim \left[\begin{pmatrix} a_{i+1} \\ f_{i+1} \end{pmatrix}, \begin{pmatrix} R_{i+1} & S_{i+1} \\ S_{i+1}^{(kl)'} & q_{i+1} \end{pmatrix}\right], \tag{5.1}$$

where

$$f_{i+1} = Z'_h a_{i+1}, \quad S_{i+1} = R_{i+1}Z_h \quad \text{and} \quad q_{i+1} = Z'_h S_{i+1}.$$

Here the prior distribution of $\lambda_{i+1(h)}$ is assumed to be a conjugate gamma distribution

(b_{i+1}, r_{i+1}). Note that b_{i+1} and r_{i+1} are estimated to be expressed in the mean and the variance of the distribution of $\log \lambda_{i+1(h)}$ such as, respectively, q_{i+1}^{-1} and $q_{i+1}^{-1}\exp(-f_{i+1})$. By integration, for $t \in I_{i+1}$,

$$P(T_h > t | T_h > \tau_i, D_i) = \left(1 + \frac{t - \tau_i}{r_{i+1}}\right)^{-b_{i+1}}$$

Thus, the survival function of individuals of the covariate vector Z_h in the time interval I_{i+1} is forecasted based on D_i by Bayes rule as follows

$$P(T_h > t | D_i) = P(T_h > t | T_h > \tau_i, D_i) P(T_h > \tau_i | D_i), \text{ for } t \in I_{i+1}, i = 1, \cdots, s-1.$$

6. Illustrations

In this section, the performance of the Bayesian estimations proposed in previous sections is illustrated, via the data in Table 1 which consist of survival times of 90 gastric carcinoma patients equally divided into two groups with respect to the type of treatments - the chemotherapy and the combination of chemotherapy and radiation therapy.

Table 1. Survival Times in days

Chemotherapy	1,63,105,129,182,216,250,262,301,301,342,354,356, 358,380,381c,383,383,388,394,408,460,489,499,524, 529c,535,562,675,676,748,748,778,786,797,945c,955 ,968,1180c,1245, 1271,1277c,1397c,1512c,1519c
Chemotherapy and Radiation	17,42,44,48,60,72,74,95,103,108,122,144, 167,170,183,185,193,195,197,208,234,235, 254,307,315,401,445,464,484,528,542,567, 577,580,795,855,882c,892c,1031c,1033c, 1306c,1335,1366,1452c, 1472c

c : censored , source : Stablein et al.(1981)

In the discount survival model the survival time of the the *j-th* individual alive at the beginning of time interval I_i assumed to have the exponential distribution with hazard rate,

$$\lambda_{i(j)} = \exp(Z'_{ij}\beta_i)$$

with $Z'_{ij} = (1, z_{ij})$ and $\beta_i = (\beta_{0i}, \beta_{1i})'$, where $z_{ij} = 0$ for the chemotherapy and $z_{ij} =$ for the combined therapy of the *j-th* individual alive at the beginning of time interval I_i. Each end point of time intervals is taken as the multiple of 30 days. We start the analysis with the initial distribution of $(\beta_0 | D_0) \sim [0, 10^2 I_2]$ reflecting lack of informations at time $t = 0$. Also with discount factors 1 and 0.75 for variances corresponding to β_0 and β_1, respectively, so that

$$B_i = \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{0.75} \end{pmatrix}$$

Figure 1 shows the considerable time variation of the mean of the parameter β_{1i} modeling the difference of the treatment effect. It changes from a positive contribution to the hazard rate in the early time to a negative contribution as time elapses and stays near zero. Figure 2 shows estimated survival functions of two treatment groups under the discount survival model. One can see that at nearly 1100 days two functions intersect, which does not agrees with estimates of survival functions under the Cox model but PL-estimates. Figure 3 shows one-step ahead forecasted survival functions under the discount survival model.

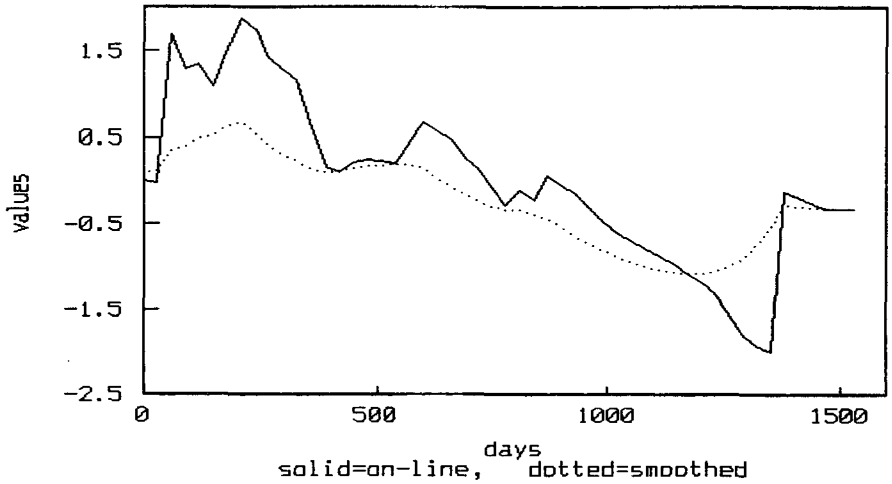

Figure 1. Estimated Mean of Parameter of Effect of Treatment Difference

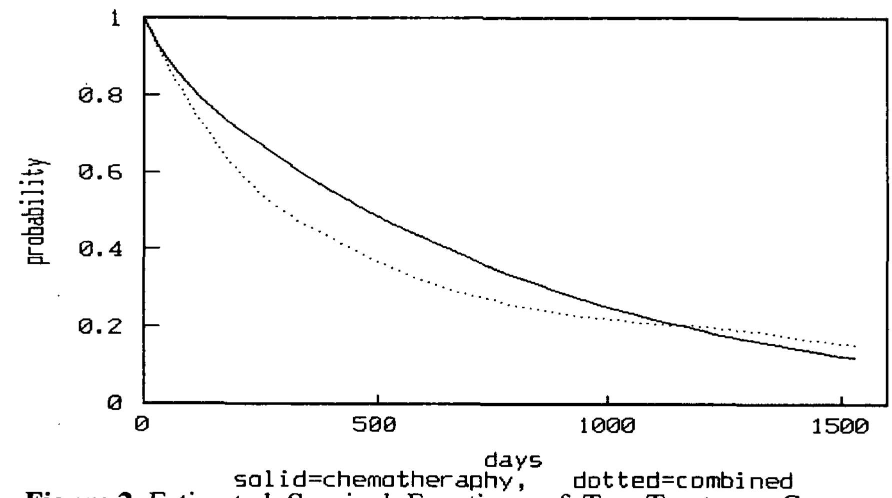

Figure 2. Estimated Survival Functions of Two Treatment Groups

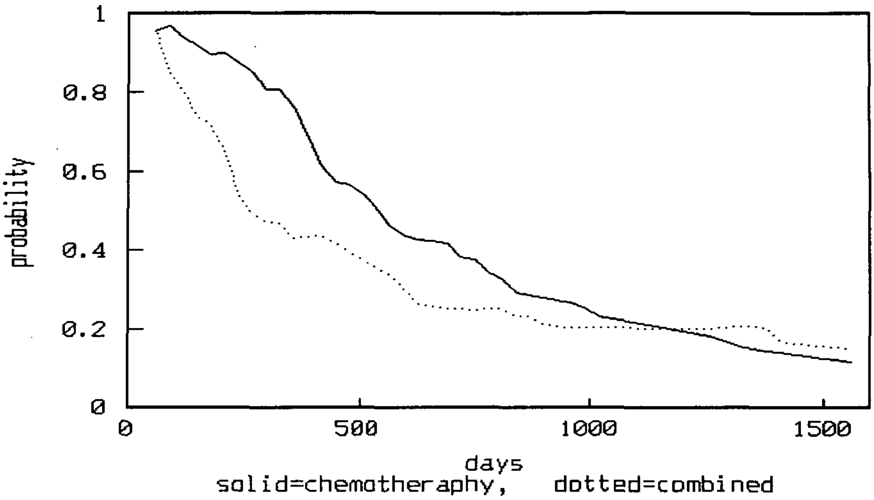

Figure 3. One-Step Ahead Forecasted Survival Functions under DSM

References

Ameen, J. R. M. and Harrison, P. J. (1985), *Normal Discount Bayesian Models*, Bayesian Statistics 2, J. M. Bernardo, M. H. DeGroot, D. V. Lindley and A. F. M. Smith (Eds.). North-Holland, Amsterdam, and Valencia University Press.

Cox, D. R. (1972), *Regression Models and Life-Tables*(with discussions), Journal of the Royal Statistical Society B, Vol. 34, 187-220.

Gamerman, D. (1991), *Dynamic Bayesian Models for Survival Data*, Applied Statistics, Vol. 40, 63-79.

Nelder, J. A. and Wedderburn, R. W. M. (1972). *Generalized linear models*. Journal of the Royal Statistical Society A, Vol. 35, 370-384.

Stablein, D. M., Carter, W. H. and Novak, J. W. (1981). *Analysis of survival data with nonproportional hazard functions*. Controlled Clinical Trials, Vol. 2, 149-159.

West, M., Harrison, P. J. and Migon, H. S. (1985). *Dynamic Generalized Linear Models and Bayesian Forecasting*, Journal of the American Statistical Associaion, Vol. 80, 73-97.