Project leuk -bayes

Maéva Caillat, Stephen Jaud, Saimourya surabhi

March 2020

1   Introduction

Several authors have discussed Bayesian inference for censored survival data where the integrated baseline hazard function is to be estimated non-parametrically Kalbfleisch (1978), Kalbfleisch and Prentice (1980), Clayton (1991), Clayton (1994). Clayton (1994) formulates the Cox model using the counting process notation introduced by Andersen and Gill (1982) and discusses estimation of the baseline hazard and regression parameters using MCMC methods. Although his approach may appear somewhat contrived, it forms the basis for extensions to random effect (frailty) models, time-dependent covariates, smoothed hazards, multiple events and so on. We show below how to implement this formulation of the Cox model in BUGS.

For subjects $i = 1...n$, we observe processes $N_i(t)$ which count the number of failures which have occurred up to time t. The corresponding intensity process $I_i(t)$ is given by:

$$I_i(t)dt = \mathbb{E}\left(dN_i(t) \mid F_{t-}\right) \tag{1}$$

where $dN_i(t)$ is the increment of N_i over the small time interval $[t, t + dt[$, and F_{t-} represents the available data just before time t. If subject i is observed to fail during this time interval, $dN_i(t)$ will take the value 1; otherwise $dN_i(t) = 0$. Hence $\mathbb{E}\left(dN_i(t) \mid F_{t-}\right)$ corresponds to the probability of subject i failing in the interval $[t, t+dt[$. As $dt \to 0$ (assuming time to be continuous) then this probability becomes the instantaneous hazard at time t for subject i . This is assumed to have the proportional hazards form:

$$I_i(t) = Y_i(t)\,\lambda_0(t)\,e^{\beta z_i} \tag{2}$$

where $Y_i(t)$ is an observed process taking the value 1 or 0 according to whether or not subject i is observed at time t and $\lambda_0(t)e^{\beta z_i}$ is the familiar Cox regression model. Thus we have observed data $D = (N_i(t), Y_i(t), z_i)_{1...n}$ and unknown parameters β and $\Lambda_0(t) = \int_0^t \lambda_0(u)du$, the latter to be estimated non-parametrically.

The joint posterior distribution for the above model is defined by

$$\mathbb{P}\left(\beta, \Lambda_0(.) \mid D\right) \sim \mathbb{P}\left(D \mid \beta, \Lambda_0(.)\right)\mathbb{P}(\beta)\mathbb{P}\left(\Lambda_0(.)\right) \tag{3}$$

For BUGS, we need to specify the form of the likelihood $\mathbb{P}\left(D \mid \beta, \Lambda_0(.)\right)$ and prior distributions for β and $\Lambda_0(.)$. Under non-informative censoring, the likelihood of the data is proportional to

$$\prod_{i=1}^{n}\left[\prod_{t \geq 0} I_i(t)^{dN_i(t)}\right] e^{-I_i(t)dt} \tag{4}$$

This is essentially as if the counting process increments $dN_i(t)$ in the time interval $[t, t + dt[$ are independent Poisson random variables with means $I_i(t)dt$:

$$dN_i(t) \sim \mathcal{P}(I_i(t)dt) \tag{5}$$

We may write:

$$I_i(t)dt = Y_i(t)\,e^{\beta z_i}\,d\Lambda_0(t) \tag{6}$$

where $d\Lambda_0(t) = \lambda_0(t)dt$ is the increment or jump in the integrated baseline hazard function occurring during the time interval $[t, t + dt[$. Since the conjugate prior for the Poisson mean is the gamma distribution, it would be convenient if $\Lambda_0(.)$ were a process in which the increments $d\Lambda_0(t)$ are distributed according to gamma distributions. We assume the conjugate independent increments prior suggested by Kalbfleisch (1978), namely

$$d\Lambda_0(t) \sim \Gamma\left(c\, d\Lambda_0^*(t), c\right) \tag{7}$$

Here, $d\Lambda_0^*(t)$ can be thought of as a prior guess at the unknown hazard function, with c representing the degree of confidence in this guess. Small values of c correspond to weak prior beliefs. In the example below, we set $d\Lambda_0^*(t) = rdt$ where r is a guess at the failure rate per unit time, and dt is the size of the time interval.

The above formulation is appropriate when genuine prior information exists concerning the underlying hazard function. Alternatively, if we wish to reproduce a Cox analysis but with, say, additional hierarchical structure, we may use the multinomial-Poisson trick described in the BUGS manual. This is equivalent to assuming independent increments in the cumulative 'non-informative' priors. This formulation is also shown below.

The fixed effect regression coefficients β are assigned a vague prior

$$\beta \sim \mathcal{N}(0, 10^6) \tag{8}$$

For further description about the project In OpenBUGS : http://www.openbugs.net/Examples/Leuk.html

2 Implementation

Calculate the posterior distribution
$\theta = (\beta, d\Lambda_0(t_1), d\Lambda_0(t_2), ..., d\Lambda_0(t_p))$
$D = \left((N_{i1}, ..., N_{ip}), (Y_{i1}, ..., Y_{ip}), z_i\right)_{1..n}$
$\pi(\beta) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{\beta}{2\sigma^2}}$ with $\sigma^2 = 10^6$
$\pi(d\Lambda_0) = \prod_{i=1}^p f_{\Gamma(c\, r\, dt_j, c)}(d\Lambda_0(t_j))$
$\pi(D|\theta) = \prod_{i=1}^n e^{-\int I_i(t)dt} \prod_{j=1}^p \left[Y_{ij}\lambda_0(t_j)e^{\beta z_i}\right]^{dN_{ij}}$

$$
\begin{aligned}
\pi(\theta|D) \;\propto\;& \pi(D|\theta)\,\pi(\theta) \\
\propto\;& \pi(D|\theta)\,\pi(\beta)\,\pi(d\Lambda_0) \\
\propto\;& \prod_{i=1}^n \left(e^{-e^{\beta z_i}\sum_{j=1}^p d\Lambda_0(t_j)} \prod_{j=1}^p \left[Y_{ij}d\Lambda_0(t_j)e^{\beta z_i}\right]^{dN_{ij}}\right) f_{\mathcal{N}(0,\sigma^2)}(\beta) \prod_{j=1}^p f_{\Gamma(c\, r\, dt_j, c)}(d\Lambda_0(t_j)) \\
\propto\;& \exp\left(-\sum_{i=1}^n e^{\beta z_i}\sum_{j=1}^p d\Lambda_0(t_j)\right) \prod_{i=1}^n \left(\prod_{j=1}^p \left[Y_{ij}d\Lambda_0(t_j)e^{\beta z_i}\right]^{dN_{ij}}\right) f_{\mathcal{N}(0,\sigma^2)}(\beta) \prod_{j=1}^p f_{\Gamma(c\, r\, dt_j, c)}(d\Lambda_0(t_j))
\end{aligned}
$$

Hence the log-density :

$$
\begin{aligned}
\log\pi(\theta|D) - C \;=\;& -\sum_{i=1}^n e^{\beta z_i}\sum_{j=1}^p d\Lambda_0(t_j) + \sum_{i=1}^n\sum_{j=1}^p dN_{ij}\log\left[Y_{ij}d\Lambda_0(t_j)e^{\beta z_i}\right] + \log f_{\mathcal{N}(0,\sigma^2)}(\beta) \\
& + \sum_{j=1}^p \log f_{\Gamma(c\, r\, dt_j, c)}(d\Lambda_0(t_j))
\end{aligned}
$$