Inference in shared frailty Cox models using \textit{R}

Christian B. Pipper

c.pipper@biostat.ku.dk
A typical example: Diabetic retinopathy data

Diabetic retinopathy results in blindness.

Goal: to examine whether or not a certain laser treatment has a positive effect in delaying onset of blindness for a person with adult diabetes.

Study design

- 83 patients with adult diabetes.
- One eye of each patient randomly chosen for treatment the other left untreated.

Observations for each eye

- Time to blindness while in study or just time spent in study.
- Person (cluster) to which the eye belongs.
- Laser treatment or no treatment.
Problem!

• Person (cluster) heterogeneity resulting in dependence between times to blindness for a patient’s two eyes.

• This needs to be reflected in an analysis of the data (Lin, Stat. Med. 1994).

Possible solution

• Model the unobserved heterogeneity (the frailty concept).
A quick introduction to the shared frailty Cox model

The ingredients

- $T_{ik}, C_{ik}, Z_{ik}$ survival time, censoring time, and censoring time of the $i$th individual in the $k$th cluster.
- Corresponding $X_{ik} = \min(T_{ik}, C_{ik})$, $N_{ik}(t) = I(X_{ik} \leq t)$, $Y_{ik}(t) = I(X_{ik} \geq t)$ event time, counting process, and at risk process.
- $F_k$ frailty corresponding to $k$th cluster, $F_k > 0$ with laplace transform $\phi_{\theta}$ for some parameter $\theta$.
- Observable filtration $\mathcal{F}_t = \bigvee_{i,k} \sigma\{N_{ik}(s), Y_{ik}(s), Z_{ik} : 0 \leq s \leq t\}$
- Modelling filtration $\mathcal{H}_t = \bigvee_{i,k} \sigma\{N_{ik}(s), Y_{ik}(s), Z_{ik}, F_k : 0 \leq s \leq t\}$
Formulating the model informally

- Assume independence between failure times given frailties and covariates
- Hazard rate/instantaneous rate of failure conditional on frailties: \( F_k \lambda_0(t) \exp(\beta^T Z_{ik}) \)

Formally

- Assume independent censoring
- Assume hazard rate as before
- Intensity

\[
\lambda_{ik}^H(t) = F_k Y_{ik}(t) \lambda_0(t) \exp(\beta^T Z_{ik}) ,
\]

meaning that

\[
M_{ik}^H(t) = N_{ik}(t) - \int_0^t \lambda_{ik}^H(s) ds
\]

is a martingale w.r.t. \( \mathcal{H}_t \)
Important to note

- The covariate effect $\beta$ needs to be interpreted as a within cluster log(relative risk) increase/decrease!

- Marginal hazard rate

$$
\lambda_0(t) \exp(\beta^T Z_{ik}) \times \frac{-D_{\phi\theta}\{\exp(\beta^T Z_{ik})\Lambda_0(t)\}}{\phi_{\theta}\{\exp(\beta^T Z_{ik})\Lambda_0(t)\}}.
$$

- Gamma frailty $\phi_{\theta}(u) = (1 + \frac{u}{\theta})^{-\theta}$, with $\theta$ form parameter and $\theta^{-1}$ scale parameter.

- Thus $RR_{population} = RR_{within} \cdot f(t)$ with $f(0) = 1$ and $f(t) \to RR_{within}^{-1}$, $t \to \infty$. 
Traditional inference using the EM-algorithm

- E-step for given $\theta$:
  \[
  \hat{F}_k = E(F_k | \mathcal{F}_\tau)
  \]

- M-step for given $\theta$: maximize
  \[
  \prod_{ik} \prod_{t \leq \tau} \{\hat{F}_k \lambda_0(t) \exp(\beta^T Z_{ik})\}^{dN_{ik}(t)} \times \exp\left\{- \sum_{i,k} \hat{F}_k \Lambda_0(X_{ik}) \exp(\beta^T Z_{ik})\right\}
  \]
  in the other parameters

- maximize the resulting profile likelihood in $\theta$ to find MLE of the parameters
Issues using the EM-algorithm and ML inference

- Double iterative... very slow...
- Only large sample results of the ML estimators for shared gamma frailties (Parner, 1998).
Inference in R using penalized likelihood

- For given $\theta$ maximize

$$\prod_{t \leq \tau} \prod_{i,k} \left\{ \frac{F_k \exp(\beta^T Z_{ik})}{\sum_{i,k} F_k Y_{ik}(t) \exp(\beta^T Z_{ik})} \right\} dN_{ik}(t) - g(F_1, \ldots, F_k, \theta)$$

- $g$ is the penalty function.

Gamma frailties

- Gamma frailty: $g(F_1, \ldots, F_k, \theta) = \frac{1}{\theta} \sum_k (\log(F_k) - F_k)$

- Same estimates as with EM for given $\theta$.

- Profile likelihood for $\theta$ is easily found from the penalized likelihood.

- For more details, see Therneau et al. (2003).
Issues using PPL

- PPL is a lot faster than EM.
- Easy to use fast running implementation in R by Therneau and Grambsch.
- SE’s of $\hat{\beta}$ are calculated for given $\hat{\theta}$ may not be correct if dependence is present (usually not the case). Easy to remedy, though.
- A penalized likelihood procedure for log-normal frailties is also implemented. However, for log-normal frailties there is no theory for the resulting estimators.
Analysis of the Diabetic Retinopathy data

Gamma frailty

> analyse<-coxph(Surv(time,cens)~trt+frailty(person.id,dist="gamma"))
> analyse

Call:
coxph(formula = Surv(time, cens) ~ trt + frailty(person.id, dist = "gamma"))

                              coef se(coef) se2  Chisq DF  p
trt                     -1.5  0.286  0.281  27.5  1.0 1.6e-07
frailty(person.id, dist =      46.6  34.4  8.0e-02

Iterations: 6 outer, 24 Newton-Raphson
        Variance of random effect= 0.773  I-likelihood = -307.4
Degrees of freedom for terms= 1.0 34.4
Likelihood ratio test=99 on 35.3 df, p=5.88e-08  n= 166
>
Home made program for getting the right confidence interval of the trt effect

```r
conf.int.beta.frail <- function(covariate, limits, precision, time, cens, frail.id) {

    imax <- round((limits[2] - limits[1]) / precision, 1) + 1
    profilelik <- rep(0, imax)
    for (i in 1:imax) {
        beta <- limits[1] + i * precision
        analyse <- coxph(Surv(time, cens) ~ offset(beta * covariate) + frailty(frail.id))
        profilelik[i] <- analyse$history$frailty$c.loglik
    }
    maxprofile <- max(profilelik)
    point <- (1:imax)[profilelik == maxprofile]
    i <- 1
    while (profilelik[i] + 3.84 / 2 < maxprofile & i < point) {
        lower <- limits[1] + precision * i
        i <- i + 1
    }
    while (profilelik[i] + 3.84 / 2 > maxprofile & i < imax) {
        upper <- limits[1] + precision * i
        i <- i + 1
    }
    plot((limits[1] + (1:imax) * precision), profilelik, type = "l", lwd = 4, xlab = "beta", ylab = "log(Profile PL)"
    abline(maxprofile - 3.84 / 2, 0, lwd = 4, col = "red")
    return(c(lower, upper))
}
```
Running the program

```r
> conf.int.beta.frail(trt,c(-2.5,-0.5),0.01,time,cens,person.id)
[1] -2.15 -0.93
>
• Compared to $(-1.5 - 1.96 \cdot 0.286; -1.5 + 1.96 \cdot 0.286) = (-2.06; -0.94)$.
```

![Figure 1: log profile partial likelihood as a function of trt effect](image)
Stratified Cox (Gross and Huber, 1987)

```r
> analyse1<-coxph(Surv(time,cens)~trt+strata(person.id))
> analyse1
Call:
coxph(formula = Surv(time, cens) ~ trt + strata(person.id))

 coef  exp(coef) se(coef)    z  p
trt  -1.55    0.213   0.348 -4.44 8.8e-06

Likelihood ratio test=26.1 on 1 df, p=3.28e-07 n= 166
>

Not accounting for person heterogeneity

```r
> analyse2<-coxph(Surv(time,cens)~trt)
> analyse2
Call:
coxph(formula = Surv(time, cens) ~ trt)

 coef  exp(coef) se(coef)    z  p
trt  -1.29    0.275   0.276 -4.67 3.1e-06

Likelihood ratio test=25.2 on 1 df, p=5.28e-07 n= 166
>
Some nice facts to know

Formal test for treatment effect in the frailty model

```r
> analyse.no.trt<-coxph(Surv(time,cens)~frailty(person.id,dist="gamma"))
> analyse.no.trt
Call:
coxph(formula = Surv(time, cens) ~ frailty(person.id, dist = "gamma"))

              coef se(coef)  se2 Chisq DF  p
frailty(person.id, dist = 0.42 0.4 0.28

Iterations: 5 outer, 14 Newton-Raphson
Variance of random effect= 0.00597   I-likelihood = -322
Degrees of freedom for terms= 0.4
Likelihood ratio test=0.84 on 0.4 df, p=0.141  n= 166
>
```

- Use I-likelihood in "analyse" and "analyse.no.trt" to make likelihood ratio test. \(-2logQ = 2 \cdot (322 - 307.4) = 29.2\) with 1 degree of freedom (p-value=6.87 \times 10^{-8}). Similar to approximate Wald test displayed in "analyse".
Formal test for $\theta = \infty$

```r
lr.test
[1] 3.942682
1 - pchisq(lr.test, 1)
[1] 0.04707567
```

- Similar - to some extent - to approximate Wald test displayed in "analyse".
References


• Lin, D.Y. (1994). Cox regression analysis of multivariate failure time data: the marginal


• Therneau, T.M., Grambsch, P.M., and Pankratz, V.S. (2003). Penalized survival models