Comparing Marginal and Random Effects (Frailty) Models

Terry M. Therneau
Mayo Clinic

April 1998

My Perspective

- Software
- Examples
- Reliability
- Areas that need work

Software Maturity

- Exists but you wouldn't want it
- Worked application(s) in a paper
  software shared with friends
- General program/ Splus function
  - available from author and/or Statlib
  - advertised
- Available in a package
- In SAS

The models attempt to deal with correlated data

- Multiple events per subject
- Multiple event types per subject
- Correlated family members
- Correlation due to sampling plan
- ⋮

Marginal models

Hueristic rationale:

Consider a weighted linear model, either the case of subject weights or a fully known variance matrix Σ.

If we fit the model ignoring the weights

- $\hat{\beta}$ usually changes very little
- $\mathrm{Var}(\hat{\beta})$ may be badly incorrect

Marginal Cox models

There are three steps

- Set up the data set appropriately.
- Fit the data as though it were independent observations.
- Fix up the variance, post fit, using a grouped jackknife.
  - Compute the *dfbeta* residual for each observation
  - Add these up to get the $D_{j.}$= effect of the jth group
  - $V = D'D$
  - (Identical to the working-independence method of GEE)

Frailty Models

A good overview of problems and motivation is given in Aalen, "Heterogeneity in survival analysis", Statistics in Medicine, 1988.

Assume that subject i from family j has an inherent, and unmeasured risk (frailty) ω_j, such that

$$\begin{aligned} \lambda_i(t) &= \lambda_0(t)\omega_j e^{X_i\beta} \\ &= \lambda_0(t)e^{X_i\beta+Z_i\gamma}\,, \end{aligned}$$

where $Z_{ij} = 1$ iff subject i is a member of family j.

Further assume

- The subjects are conditionally independent given γ.
- γ is a random variable from a known distribution.

then a solution can be computed.

Issues

Different models available

- Distribution of the frailty
  - gamma
  - log-normal
  - positive stable
  - inverse Gaussian
  - . . . Does it make any difference?
- Pattern of random effects
  - 1 per subject or family
  - Nested $\rightarrow$ block diagonal variance for γ
  - Crossed

Each combination has it's own, unique variant of the EM algorithm.

Penalized models

PPL = Penalized Partial Likelihood = PL - penalty

Formally, let the Cox model hazard be

$$\lambda_i(t) = \lambda_0(t) e^{X_i\beta + Z_i\gamma}$$

where

- β = unpenalized effects
- γ = penalized effects

So

$$\text{PPL} = \text{PL}(\beta, \gamma; \text{data}) - p(\gamma, \theta)$$

A simple example is "ridge regression" with

$$p(\gamma, \theta) = \theta \sum \gamma_j^2.$$

Solution

If the ancillary parameter(s) θ is known

- The solution of the Cox model for (β, γ) is simple.
  - Straightforward addition to a standard Cox program.
  - Need p, p' and p''
- Well-recognized variance formulas.
  - H^{-1}, where $H_{ij} = \partial^2 \mathrm{PLL}/\partial\beta_i\partial\beta_j$
  - $H^{-1}\mathcal{I}H^{-1}$, where $\mathcal{I}_{ij} = \partial^2 \mathrm{PL}/\partial\beta_i\partial\beta_j$
- Definitions for degrees of freedom.
- Connections to other techniques such as GCV, AIC.
- An array of numerical techniques available.

Implementation

This has been added to the S-Plus `coxph` function.

A given method requires

- define the penalty p, and its derivatives p' and p''
- define the logic for determining θ
- setup

These can be user-defined functions.

The gamma frailty (Nielson et al 1992) and log-normal frailty (McGilchrist 93) map exactly onto this structure.

See ftp://ftp.mayo.edu/pub/therneau

Rat data

- Female rats from Mantel et. al.
- 3 rats/litter, one treated + two placebo
- 50 litters

Marginal analysis

- The events are unordered: the time scale for each observation starts at zero and goes to event/censoring.
- The event types (death of rat A, death of rat B) are identical: only one stratum.
- The data set will have 150 observations, identical to an "iid" data set.
- Variables `time, status, rx` and `litter`.

```
> coxph(Surv(time, status) ~ rx + cluster(litter),
                        data=rats)

     coef   exp(coef) se(coef)  robust se    z      p
rx  0.898      2.46    0.317       0.3     2.99 0.0028

Likelihood ratio test=7.87  on 1 df, p=0.00503  n= 150
```

Gamma frailty

```
> coxph(Surv(time, status) ~ rx + gfrail(litter), rats)
    coef se(coef)   se2 Chisq   DF     p
rx 0.906  0.323    .319  7.88  1.0 0.005

Iterations: 6 outer, 19 Newton-Raphson
         Variance of random effect= 0.474
Degrees of freedom for terms=  1.0 13.8
Likelihood ratio test=9.4  on 14.82 df, p=0.847  n= 150
```

Gaussian (normal) frailty

```
> coxph(Surv(time, status) ~ rx + nfrail(litter), rats)
    coef se(coef)   se2 Chisq   DF     p
rx 0.904 0.322    0.318  7.89  1.0 0.005

Iterations: 5 outer, 14 Newton-Raphson
         Variance of random effect= 0.39
Degrees of freedom for terms=  1.0 11.4
Likelihood ratio test=34  on 12.39 df, p=0.000864  n= 150
```

Colon cancer data

- 929 patients
  - 315 Observation
  - 310 Levamisole
  - 304 Levamisole + 5FU
- Time to death and progression for each subject
  - 423 No events
  - 92 One event
  - 414 Two events
- Up to 9 years of follow-up

Marginal fit

```
> coxph(Surv(time, status) ~ rx + extent + node4 +
                              cluster(id) + strata(etype),
                                data=colon)

              coef exp(coef) se(coef) robust se      z       p
    rxLev  -0.0362     0.964   0.0768    0.1056 -0.343 7.3e-01
rxLev+5FU -0.4488     0.638   0.0840    0.1168 -3.842 1.2e-04
   extent  0.5155     1.674   0.0796    0.1097  4.701 2.6e-06
    node4  0.8799     2.411   0.0681    0.0961  9.160 0.0e+00
```

	WLW	Gamma $\sigma = 1$	Gamma Frailty	Normal Frailty
Lev vs Obs	-0.04	-0.04	0.04	-0.03
Lev/5FU vs Obs	-0.45	-0.57	-0.51	-0.79
Extent	0.52	0.81	1.34	1.13
Nodes > 4	0.88	1.48	2.33	2.12
σ^2	1	–	8.06	6.95
iterations	4	1/5	10/76	8/103
'df'	4	513	1377	931

Hidden Covariate Example

- True hazard for a subject is $\lambda(t) = \exp(x_1 - x_2)$.
- X_1 is uniform(-2,2) and unknown.
- X_2 is the 0/1 treatment variable, known.
- All observations have the same total follow-up.
- Multiple sequential events per subject.
- We purposely have chosen that the hidden variate have a larger effect than treatment.
- The true coefficients are +1 and -1.
- 250 replications, 100 subjects per arm

	Number of events			
	0	1-2	3-5	6-10
Control	11	21	34	34
Treatment	29	46	26	1

(606 events)

Fit two marginal models:

- Events are sequential, so time is (0, first event], (first, second], ...
- A subject is not at risk for a second event until they have had their first.
- Conditional model: First event, second event, etc are each in a separate strata. (Events may change the subject).
- AG model: All observations in one strata. (Poisson process).

```
fit1 <- coxph(Surv(start, stop, status) ~
                                    rx + strata(enum) + cluster(
fit2 <- coxph(Surv(start, stop, status) ~ rx + cluster(id))

fit3 <- coxph(Surv(start, stop, status) ~
                                    rx + strata(enum) + gfrail(i
```

Summary

- Seems to work.
- The precision of the random effect is poor.
- The connection to degrees of freedom is unclear.
- Less flexible than marginal models.
- Not yet reliable.

Open Issues

- Code
  - Beta version available on ftp://ftp.mayo.edu/pub/therneau
  - Will be a part of Splus proper eventually.
- Variance
  - Which estimate is better?
  - Should we bootstrap it?
- Frailty
  - What other distributions can be accomodated?
  - Crossed or nested random effects.
  - How well do the more general, extensible rules work? AIC, BIC, PRESS, CIC, ...