Models In Epidemiology And Biostatistics
Gordon Hilton Fick

Conditional Time To Event Models

Now let us suppose we have n subjects and we have n_i observations on subject i. We are interested in the time to an event that can recur. [not mortality, for example] For the jth observation on the ith subject, we might consider the addition of a subject specific u_i . As with our previous discussions of conditional models, we can consider clusters, groups or blocks in lieu of subjects. [then mortality might be the event]

$$\log h_{ij}(t)=\log h_0(t)+\beta_1 x_{1ij}+\beta_2 x_{2ij}+\ldots\beta_k x_{kij}+u_i$$

or:

$$h_{ij}(t)=h_0(t)e^{\beta_1 x_{1ij}+\beta_2 x_{2ij}+\ldots\beta_k x_{kij}}v_i$$

where $v_i=\exp(u_i)$ and $u_i=\log(v_i)$.

The v_i have come to be called the frailties and the u_i then the log frailties.

The frailty concept can be traced back to Greenwood and Yule (1920) on "accident proneness''. The term frailty itself was introduced by Vaupel et al. (1979) and the model was substantially promoted in a seminal paper by Clayton (1978) (without using the word "frailty") on chronic disease incidence in families.

They are called 'shared' frailties in that each subject shares the same frailty. [There are 'unshared' frailties not for discussion here]

As we considered previously with other conditional models, we can assume the outcomes [the times-to-event] are conditionally independent, given the frailty. This approach can be used for times-to-event of related individuals like family members or recurrent observations on the same person so that the times-to-event are not marginally independent.

The frailties are often assumed to come from a Gamma distribution with mean 1 and variance θ . We have not seen this family of distributions before. The file gamma.do provides some illustrations. Notice that if $\theta=0$ then we have a [regular] Cox model.

Consider a study of kidney dialysis patients. The outcome here is time to infection from time of catheter insertion. Each patient contributed 2 occasions of catheter insertion.

```
use catheter.dta
stcox age female,shared(patient) effects(lf) nohr

Cox regression --
         Breslow method for ties                Number of obs      =        76
         Gamma shared frailty                   Number of groups   =        38
Group variable: patient

No. of subjects =           76                  Obs per group: min =         2
```

```
No. of failures =           58                              avg =         2
Time at risk    =         7424                              max =         2

                                                Wald chi2(2)       =     11.66
Log likelihood  =   -181.97453                  Prob > chi2        =    0.0029
```

_t	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
age	.0061825	.012022	0.51	0.607	-.0173801	.0297451
female	-1.575675	.4626528	-3.41	0.001	-2.482458	-.6688924
theta	.4754497	.2673108				

```
Likelihood-ratio test of theta=0: chibar2(01) =     6.27 Prob>=chibar2 = 0.006

Note: Standard errors of regression parameters are conditional on theta.

. list patient lf  [the log frailties]
```

	patient	lf										
1.	1	.40450061	19.	10	-.72858614	39.	20	.07915559	59.	30	.25341003
2.	1	.40450061	20.	10	-.72858614	40.	20	.07915559	60.	30	.25341003
3.	2	.28607212	21.	11	.5163713	41.	21	-2.4487067	61.	31	.35432195
4.	2	.28607212	22.	11	.5163713	42.	21	-2.4487067	62.	31	.35432195
5.	3	.09900583	23.	12	.02425208	43.	22	-.52732037	63.	32	.18576114
6.	3	.09900583	24.	12	.02425208	44.	22	-.52732037	64.	32	.18576114
7.	4	-.70276712	25.	13	.28529655	45.	23	.36766283	65.	33	.12834751
8.	4	-.70276712	26.	13	.28529655	46.	23	.36766283	66.	33	.12834751
9.	5	.14780028	27.	14	-.54148495	47.	24	.05429806	67.	34	-.19393037
10.	5	.14780028	28.	14	-.54148495	48.	24	.05429806	68.	34	-.19393037
11.	6	-.00515658	29.	15	-.83608268	49.	25	-.07760733	69.	35	.3797819
12.	6	-.00515658	30.	15	-.83608268	50.	25	-.07760733	70.	35	.3797819
13.	7	.51871587	31.	16	.11057934	51.	26	-.47970869	71.	36	-.2434041
14.	7	.51871587	32.	16	.11057934	52.	26	-.47970869	72.	36	-.2434041
15.	8	-.5756332	33.	17	-.25467225	53.	27	.06488521	73.	37	.11519003
16.	8	-.5756332	34.	17	-.25467225	54.	27	.06488521	74.	37	.11519003
17.	9	-.15290952	35.	18	-.25982922	55.	28	.38768077	75.	38	-.51796025
18.	9	-.15290952	36.	18	-.25982922	56.	28	.38768077	76.	38	-.51796025
			37.	19	-.53238607	57.	29	.30102246			
			38.	19	-.53238607	58.	29	.30102246			

```
predict lhr,xb
twoway (line lhr age if female==0)(line lhr age if female==1),legend(off)
stcurve, cumhaz at1(female=0) at2(female=1) yscale(log)
stcurve, haz at1(female=0) at2(female=1) yscale(log)
estat phtest,detail
estat phtest, plot(female)
estat phtest, log plot(female)
estat phtest, log detail
bysort patient: replace lf=. if _n>1
graph box lf, by(female) marker(1, mlabel(patient))
```

To my knowledge, there are no marginal models available currently. A 'variance corrected' method is available in Stata.

```
stcox age female, vce(cluster patient) nohr

Cox regression -- Breslow method for ties

No. of subjects      =           76                 Number of obs   =        76
No. of failures      =           58
Time at risk         =         7424
                                                    Wald chi2(2)    =      2.74
Log pseudolikelihood =   -185.10993                 Prob > chi2     =    0.2540

                              (Std. Err. adjusted for 38 clusters in patient)
------------------------------------------------------------------------------
             |               Robust
          _t |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         age |   .0022426   .0078139     0.29   0.774    -.0130724    .0175575
      female |  -.7986869    .487274    -1.64   0.101    -1.753726    .1563526
------------------------------------------------------------------------------

.
```

The method of shared frailty is far more widely seen than the VC methods.

For time to event analysis with discrete time, we can use all of the conditional models appropriate for logistic regression or cloglog models.