Cox Model With Time Varying Covariates

Andy Grogan-Kaylor

24 Nov 2023

Introduction

The Cox Proportional Hazards Model is an important model in *event history and survival analysis*. One important aspect of the Cox Model is its ability to include *time varying covariates*, covariates whose value changes over time.

The example below draws heavily from–but is slightly adapted from–the Stata `help stcox` file.

Get Data

```
. use https://www.stata-press.com/data/r17/drugtr2, clear // simulated drug data
```

Per the Stata documentation:

> "Consider a dataset consisting of 45 observations on recovery time from walking pneumonia. Recovery time (in days) is recorded in the variable time, and there are measurements on the covariates age, drug1, and drug2, where drug1 and drug2 interact a choice of treatment with initial dosage level. The study was terminated after 30 days, so those who had not recovered by that time were censored (cured = 0)."

`stset` The Data

```
. stset time, failure(cured) // set up data for survival analysis

Survival-time data settings

         Failure event: cured!=0 & cured<.
Observed time interval: (0, time]
     Exit on or before: failure

--------------------------------------------------------------------------
         45  total observations
          0  exclusions
--------------------------------------------------------------------------
         45  observations remaining, representing
         36  failures in single-record/single-failure data
      677.9  total analysis time at risk and under observation
                                                At risk from t =         0
                                     Earliest observed entry t =         0
                                          Last observed exit t =        30
```

Model 1: Drugs Are *Time Invariant* Covariates

```
. stcox age drug1 drug2 // Cox model

        Failure _d: cured
  Analysis time _t: time

Iteration 0:  Log likelihood = -116.54385
Iteration 1:  Log likelihood = -102.77311
```

```
Iteration 2:  Log likelihood = -101.92794
Iteration 3:  Log likelihood = -101.92504
Iteration 4:  Log likelihood = -101.92504
Refining estimates:
Iteration 0:  Log likelihood = -101.92504

Cox regression with Breslow method for ties

No. of subjects =    45                              Number of obs =     45
No. of failures =    36
Time at risk    = 677.9
                                                     LR chi2(3)    =  29.24
Log likelihood = -101.92504                          Prob > chi2   = 0.0000
```

_t	Haz. ratio	Std. err.	z	P>|z|	[95% conf.	interval]
age	.8759449	.0253259	-4.58	0.000	.8276873	.9270162
drug1	1.008482	.0043249	1.97	0.049	1.000041	1.016994
drug2	1.00189	.0047971	0.39	0.693	.9925323	1.011337

```
. est store M1 // store estimates
```

Model 2: Drugs Are *Time Varying Covariates*

```
. stcox age, tvc(drug1 drug2) // Cox model
        Failure _d: cured
  Analysis time _t: time

Iteration 0:  Log likelihood = -116.54385
Iteration 1:  Log likelihood = -104.50191
Iteration 2:  Log likelihood = -103.87961
Iteration 3:  Log likelihood = -103.87525
Iteration 4:  Log likelihood = -103.87525
Refining estimates:
Iteration 0:  Log likelihood = -103.87525

Cox regression with Breslow method for ties

No. of subjects =    45                              Number of obs =     45
No. of failures =    36
Time at risk    = 677.9
                                                     LR chi2(3)    =  25.34
Log likelihood = -103.87525                          Prob > chi2   = 0.0000
```

_t	Haz. ratio	Std. err.	z	P>|z|	[95% conf.	interval]
main						
age	.8786593	.0250789	-4.53	0.000	.8308552	.9292139
tvc						
drug1	1.000272	.000335	0.81	0.416	.9996161	1.000929
drug2	.9998618	.000364	-0.38	0.704	.9991486	1.000576

```
Note: Variables in tvc equation interacted with _t.

. est store M2 // store estimates
```

Model 3: Drugs Are *Time Varying Covariates* (*Manually Specified*)

```
. generate id=_n // multiple record data needs an id

. streset, id(id) // `streset` the data
-> stset time, id(id) failure(cured)

Survival-time data settings

           ID variable: id
         Failure event: cured!=0 & cured<.
Observed time interval: (time[_n-1], time]
     Exit on or before: failure
```

```
------------------------------------------------------------------------------
         45  total observations
          0  exclusions
------------------------------------------------------------------------------
         45  observations remaining, representing
         45  subjects
         36  failures in single-failure-per-subject data
      677.9  total analysis time at risk and under observation
                                                At risk from t =         0
                                     Earliest observed entry t =         0
                                          Last observed exit t =        30

. stsplit, at(failures) // split data at each recovery time
(31 failure times)
(812 observations (episodes) created)

. generate drug1emt = drug1 * _t // manual interaction of drug1 and time

. generate drug2emt = drug2 * _t // manual interaction of drug2 and time

. stcox age drug1emt drug2emt // Cox model

        Failure _d: cured
  Analysis time _t: time
       ID variable: id

Iteration 0:  Log likelihood = -116.54385
Iteration 1:  Log likelihood = -104.50191
Iteration 2:  Log likelihood = -103.87961
Iteration 3:  Log likelihood = -103.87525
Iteration 4:  Log likelihood = -103.87525
Refining estimates:
Iteration 0:  Log likelihood = -103.87525

Cox regression with Breslow method for ties

No. of subjects =    45                                 Number of obs =    857
No. of failures =    36
Time at risk    = 677.9
                                                        LR chi2(3)    =  25.34
Log likelihood = -103.87525                             Prob > chi2   = 0.0000

------------------------------------------------------------------------------
          _t | Haz. ratio   Std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
         age |   .8786593   .0250789    -4.53   0.000     .8308552    .9292139
    drug1emt |   1.000272    .000335     0.81   0.416     .9996161    1.000929
    drug2emt |   .9998618    .000364    -0.38   0.704     .9991486    1.000576
------------------------------------------------------------------------------

. est store M3 // store estimates
```

Nice Table of Estimates to Compare Models

```
. est table M1 M2 M3, star equations(1)

--------------------------------------------------------------
    Variable |       M1              M2              M3
-------------+------------------------------------------------
#1           |
         age | -.13245204***   -.12935802***   -.12935802***
       drug1 |  .00844606*
       drug2 |  .00188866
    drug1emt |                                    .0002724
    drug2emt |                                 -.00013819
-------------+------------------------------------------------
tvc          |
       drug1 |                   .0002724
       drug2 |                -.00013819
--------------------------------------------------------------
                        Legend: * p<0.05; ** p<0.01; *** p<0.001
```