Stata Lecture Notes
Class 7

The following detail the survival analysis example from the Class 7 lecture notes. These notes make use of the dataset found in the Stata file `nepal_class7.dta`

1. Check the codebook for the key variables in the analysis:

```
. codebook stime cens parity nblind gestage treat male

stime ------------------------------------------------------------- (unlabeled)

                  type:  numeric (float)

                 range:  [1,180]                      units:  1
         unique values:  164                      missing .:  0/10,295

                  mean:   158.108
              std. dev:   49.1441

           percentiles:        10%       25%       50%       75%       90%
                                89       171       174       180       180

cens ------------------------------------------------------------- (unlabeled)


                  type:  numeric (float)

                 range:  [0,1]                        units:  1
         unique values:  2                        missing .:  0/10,295

            tabulation:  Freq.  Value
                         9,651  0
                           644  1

parity  ----------------------------------------------------------- (unlabeled)

                  type:  numeric (float)

                 range:  [0,15]                       units:  1
         unique values:  15                       missing .:  256/10,295

                  mean:   2.32563
              std. dev:   2.14178

           percentiles:        10%       25%       50%       75%       90%
                                 0         1         2         3         5

nblind  ----------------------------------------------------------- (unlabeled)


                  type:  numeric (float)
                 label:  nb

                 range:  [0,1]                        units:  1
         unique values:  2                        missing .:  0/10,295

            tabulation:  Freq.   Numeric  Label
                         9,372         0  No night blind
                           923         1  Night blind
```

```
gestage  --------------------------------------------------------- (unlabeled)

                  type:  numeric (float)

                 range:  [28,46]                      units:  1
         unique values:  19                       missing .:  612/10,295

                  mean:    38.017
              std. dev:    3.7041

           percentiles:        10%       25%       50%       75%       90%
                                33        36        38        40        42

treat  ----------------------------------------------------------- (unlabeled)

                  type:  numeric (float)
                 label:  allocl

                 range:  [1,3]                        units:  1
         unique values:  3                        missing .:  0/10,295

            tabulation:  Freq.   Numeric  Label
                         3,265         1  Beta C
                         3,387         2  Placebo
                         3,643         3  Vit A

male  ------------------------------------------------------------- (unlabeled)

                  type:  numeric (float)
                 label:  sex

                 range:  [0,9]                        units:  1
         unique values:  3                        missing .:  0/10,295

            tabulation:  Freq.   Numeric  Label
                         4,966         0  female
                         5,195         1  male
                           134         9  Missing
```

A couple of things to notice:

- Look at the codebook entry for `stime`: We have already modified the survival time variable by adding 1 to all values to account for the survival times of 0 days. We have also already censored all observations at 180 days. So the range for this variable is now [1, 180].
- Notice the missing observations for `parity` and `gestage`.

2. Create the "survival" data set by defining it in Stata:

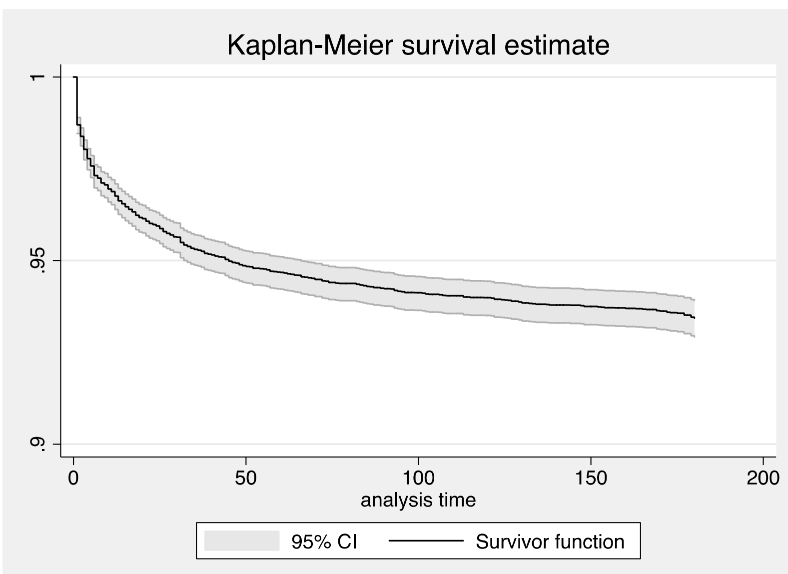
Here we define `stime` as the "time to event" variable and define "`cens=1`" to be an event/failure. This matches our data where `cens=1` means the infant died and `cens=0` means the observation was censored.

```
. stset stime, failure(cens=1)

     failure event:  cens == 1
obs. time interval:  (0, stime]
 exit on or before:  failure

------------------------------------------------------------------------------
     10,295  total observations
          0  exclusions
------------------------------------------------------------------------------
     10,295  observations remaining, representing
        644  failures in single-record/single-failure data
  1,627,725  total analysis time at risk and under observation
                                                at risk from t =         0
                                     earliest observed entry t =         0
                                          last observed exit t =       180
```

Since we have already modified survival time to change the 0 day survivals to 1 day survivals, we can see there are no exclusions (compare to slide 12 of Class Notes!)

3. Next look at an overall view of infant survival with 95% confidence bands:

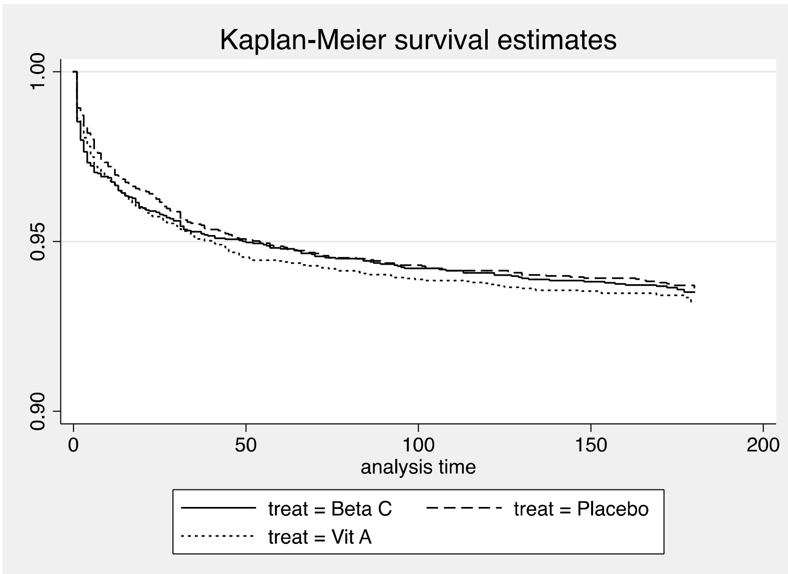
```
. sts graph, ylab(.9 (.05) 1)

         failure _d:  cens == 1
   analysis time _t:  stime
```



Notice that the survival curve is steepest in the first 20 days and that by the end of the observed time (180 days) overall survival is above 90%.

3. We can also look at infant survival by treatment group:

```
. sts graph, by(treat) ylab(.9 (.05) 1)

         failure _d:  cens == 1
   analysis time _t:  stime
```



There doesn't appear to be much difference in the survival curves for the three treatment groups. The Vitamin A group appears to have slightly lower survival than the other two treatment groups.

```
. stsum, by(treat)

         failure _d:  cens == 1
   analysis time _t:  stime

         |               incidence       no. of    |------ Survival time -----|
treat    | time at risk     rate        subjects        25%       50%       75%
---------+---------------------------------------------------------------------
  Beta C |       516692   .0003929          3265          .         .         .
 Placebo |       532438    .000385          3387          .         .         .
   Vit A |       578595   .0004079          3643          .         .         .
---------+---------------------------------------------------------------------
   total |      1627725   .0003956         10295          .         .         .
```

The incidence rates for the three groups are similar, with the highest incidence of death in the Vitamin A group. We can estimate the hazard ratios between the treatment groups using Cox regression:

```
. stcox i.treat

         failure _d:  cens == 1
   analysis time _t:  stime

Iteration 0:   log likelihood = -5902.8126
Iteration 1:   log likelihood = -5902.5964
Iteration 2:   log likelihood = -5902.5964
Refining estimates:
```

```
Iteration 0:   log likelihood = -5902.5964

Cox regression -- Breslow method for ties

No. of subjects =       10,295                  Number of obs    =     10,295
No. of failures =          644
Time at risk    =      1627725
                                                LR chi2(2)       =       0.43
Log likelihood  =   -5902.5964                  Prob > chi2      =     0.8056

------------------------------------------------------------------------------
          _t | Haz. Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       treat |
    Placebo  |   .9770902   .0967476    -0.23   0.815     .8047334    1.186362
      Vit A  |   1.039412   .0994985     0.40   0.686     .8616002     1.25392
------------------------------------------------------------------------------
```

There is no statistically significant association between treatment and hazard of infant death.

4. Now we will consider survival for each of the other covariates individually:

- **Gestational age (gestage):**

First create categories for gestational age: < 36 weeks, 36-38 weeks, 38-39 weeks, 39-41 weeks, and 41+ weeks

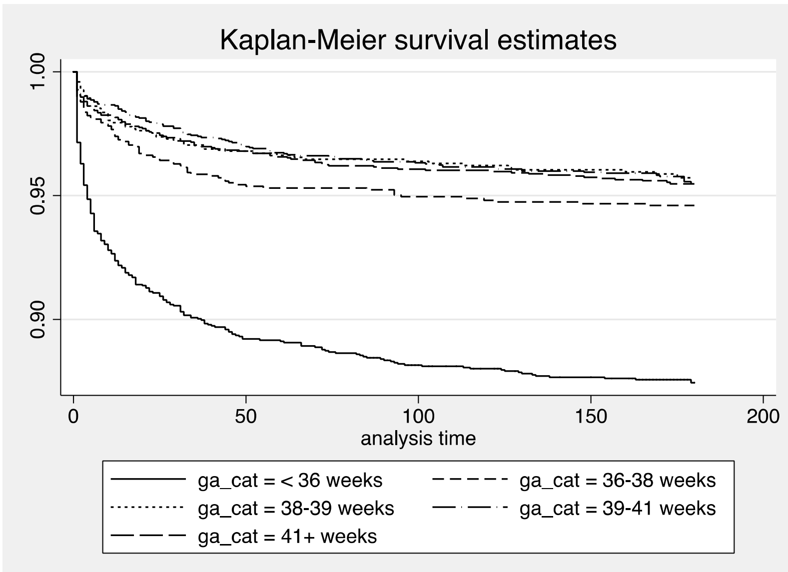
```
. gen ga_cat=1 if gestage < 36 & gestage ~=.
(8,119 missing values generated)

. replace ga_cat=2 if gestage >=36 & gestage < 38 & gestage ~=.
(1,493 real changes made)

. replace ga_cat=3 if gestage >=38 & gestage < 39 & gestage ~=.
(1,245 real changes made)

. replace ga_cat=4 if gestage >=39 & gestage < 41 & gestage ~=.
(2,499 real changes made)

. replace ga_cat=5 if gestage >=41 & gestage ~=.
(2,270 real changes made)

. label define gestcats 1 "< 36 weeks" 2 "36-38 weeks" 3 "38-39 weeks" 4 "39-41
weeks" 5 "41+
>  weeks"

. label values ga_cat gestcats
```

Look at the distribution of the sample in each of these categories:

```
. tab ga_cat

     ga_cat |      Freq.     Percent        Cum.
------------+-----------------------------------
 < 36 weeks |      2,176       22.47       22.47
36-38 weeks |      1,493       15.42       37.89
38-39 weeks |      1,245       12.86       50.75
39-41 weeks |      2,499       25.81       76.56
  41+ weeks |      2,270       23.44      100.00
------------+-----------------------------------
      Total |      9,683      100.00
```

Consider survival by gestational age category:

```
. sts graph, by(ga_cat) ylab(.9 (.05) 1)

         failure _d:  cens == 1
   analysis time _t:  stime
```



There does appear to be a difference in survival across the gestational age categories. The group with the lowest survival is the group of infants with the youngest gestational age (born the earliest). The second youngest group (36-38 weeks) has the next lowest survival. The other three groups are more difficult to distinguish.

We see the same pattern in the incidence rates below:

```
. stsum, by(ga_cat)

         failure _d:  cens == 1
   analysis time _t:  stime

         |               incidence       no. of    |------ Survival time -----|
ga_cat   | time at risk     rate        subjects        25%       50%       75%
---------+---------------------------------------------------------------------
< 36 wee |       323590   .0008128          2176          .         .         .
36-38 we |       238819   .0003266          1493          .         .         .
38-39 we |       203632   .0002554          1245          .         .         .
39-41 we |       409594   .0002564          2499          .         .         .
41+ week |       367150   .0002696          2270          .         .         .
---------+---------------------------------------------------------------------
   total |      1542785    .000387          9683          .         .         .
```

We can also look to see whether the mothers/infants with missing data are different than those without missing data. To do this, we include "Missing" as a category of gestational age and see what the incidence rate is in this group:

```
. gen ga_cat_miss=ga_cat
(612 missing values generated)

. replace ga_cat_miss=6 if ga_cat==.
(612 real changes made)

. label define gestcats_miss 1 "< 36 weeks" 2 "36-38 weeks" 3 "38-39 weeks" 4
"39-41 weeks" 5 "41+ weeks" 6 "Missing"

. label values ga_cat_miss gestcats_miss
. stsum, by(ga_cat_miss)

         failure _d:  cens == 1
   analysis time _t:  stime

         |               incidence       no. of    |------ Survival time -----|
ga_cat~s | time at risk     rate        subjects        25%       50%       75%
---------+---------------------------------------------------------------------
< 36 wee |       323590   .0008128         2176          .         .         .
36-38 we |       238819   .0003266         1493          .         .         .
38-39 we |       203632   .0002554         1245          .         .         .
39-41 we |       409594   .0002564         2499          .         .         .
41+ week |       367150   .0002696         2270          .         .         .
 Missing |        84940   .0005533          612          .         .         .
---------+---------------------------------------------------------------------
   total |      1627725   .0003956        10295          .         .         .
```

The group with missing gestational age has a high mortality compared to the other gestational age groups!

- **Parity (parity):**

Again, create categories for parity: 0, 1, 2-4, 5-8, and 9+

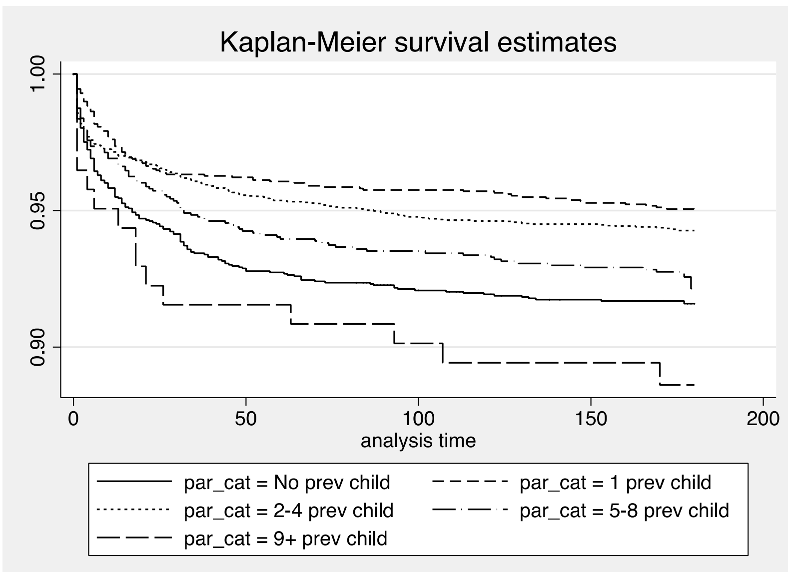
```
. gen par_cat=0 if parity==0
(8,041 missing values generated)

. replace par_cat=1 if parity==1
(2,018 real changes made)

. replace par_cat=2 if parity >=2 & parity <=4
(4,262 real changes made)

. replace par_cat=3 if parity >=5 & parity <=8
(1,363 real changes made)

. replace par_cat=4 if parity >8 & parity ~=.
(142 real changes made)

. label define par 0 "No prev child" 1 "1 prev child" 2 "2-4 prev child" 3 "5-8
prev child" 4 "9+ prev child"

. label values par_cat par
```

Look at the distribution of the sample in each of these categories:

```
. tab par_cat
       par_cat |      Freq.     Percent        Cum.
---------------+-----------------------------------
 No prev child |      2,254       22.45       22.45
  1 prev child |      2,018       20.10       42.55
2-4 prev child |      4,262       42.45       85.01
5-8 prev child |      1,363       13.58       98.59
 9+ prev child |        142        1.41      100.00
---------------+-----------------------------------
         Total |     10,039      100.00
```

Consider survival by parity category:

```
. sts graph, by(par_cat) ylab (.9 (.05) 1)
         failure _d:  cens == 1
   analysis time _t:  stime
```



There does appear to be a difference in survival across the parity categories. The group with the lowest survival is the group with the highest parity (9+ previous children). The group with no previous children (parity = 0) had the next lowest survival. Then the group with 5-8 previous children, then 2-4 previous children, and the highest survival was the group with 1 previous child.

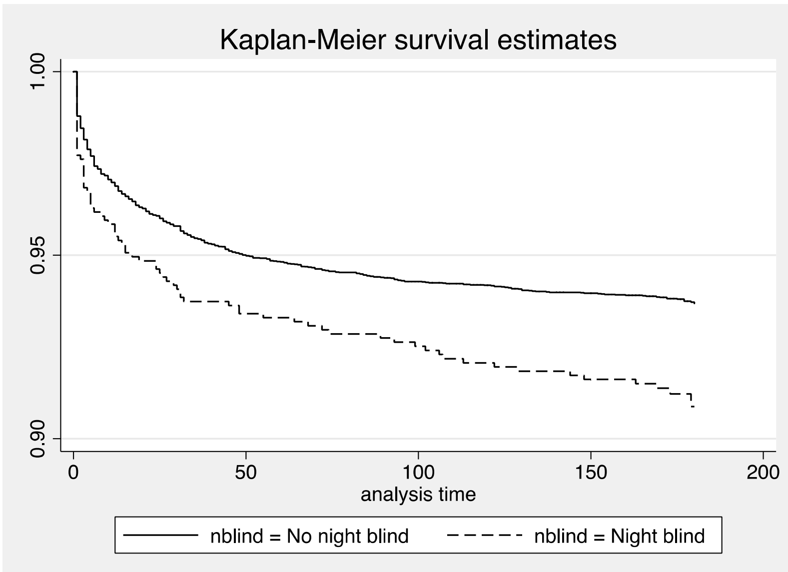
We see the same pattern in the incidence rates below:

```
. stsum, by(par_cat)

         failure _d:  cens == 1
   analysis time _t:  stime

         |               incidence       no. of    |------ Survival time -----|
par_cat  | time at risk     rate        subjects        25%       50%       75%
---------+---------------------------------------------------------------------
 No prev |       343459   .0005241          2254          .         .         .
1 prev c |       323928   .0002994          2018          .         .         .
2-4 prev |       689362   .0003452          4262          .         .         .
5-8 prev |       221722   .0004555          1363          .         .         .
 9+ prev |        22507   .0007109           142          .         .         .
---------+---------------------------------------------------------------------
   total |      1600978   .0003948         10039          .         .         .
```

- **Night blindness (nblind):**

Consider survival by whether or not the mother was night blind:

```
. sts graph, by(nblind) ylab(.9 (.05) 1)

         failure _d:  cens == 1
   analysis time _t:  stime
```



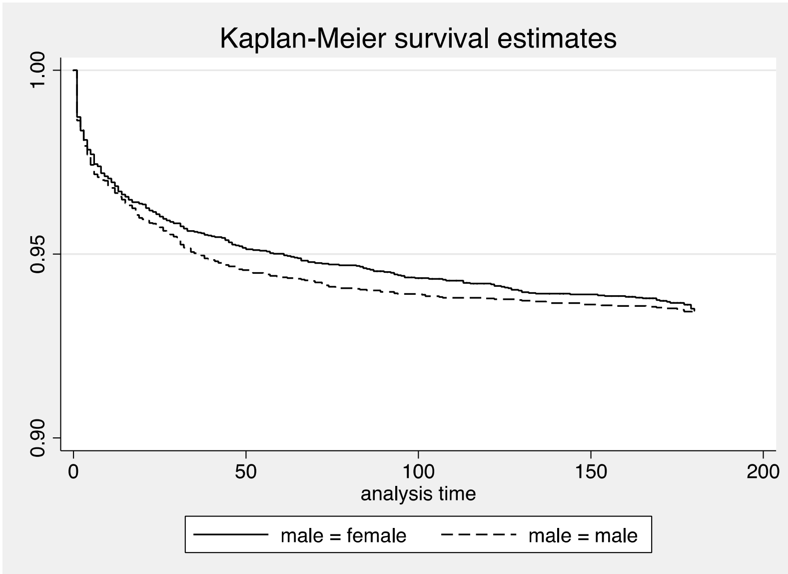
Survival is highest in the group of infants whose mothers were not night blind. We see the same result in the incidence rates below; the infants whose mothers were not night blind had a lower mortality rate:

```
. stsum, by(nblind)

         failure _d:  cens == 1
   analysis time _t:  stime

         |               incidence       no. of    |------ Survival time -----|
nblind   | time at risk     rate        subjects        25%       50%       75%
---------+---------------------------------------------------------------------
No night |      1482249   .0003805          9372          .         .         .
Night bl |       145476   .0005499           923          .         .         .
---------+---------------------------------------------------------------------
   total |      1627725   .0003956         10295          .         .         .
```

- **Gender (male):**

Consider survival by whether or not the infant was male. We indicate “`if male ~=9`” in our analysis to exclude the missing values!

```
. sts graph if male ~=9, by(male) ylab(.9 (.05) 1)

         failure _d:  cens == 1
   analysis time _t:  stime
```



The survival curves are very similar, with a slightly higher survival in the group of male infants. We see a slightly higher mortality rate for the males as well:

```
. stsum if male ~=9, by(male)

         failure _d:  cens == 1
   analysis time _t:  stime

         |               incidence       no. of    |------ Survival time -----|
male     | time at risk     rate        subjects        25%       50%       75%
---------+---------------------------------------------------------------------
  female |       797431   .0003887          4966          .         .         .
    male |       829462   .0003991          5195          .         .         .
---------+---------------------------------------------------------------------
   total |      1626893    .000394         10161          .         .         .
```

5. Now we build a Cox proportional hazards model for the hazard of infant death.

We will consider different characterizations of the gestational age and parity variables before finally fitting a multivariable Cox regression model:

- **Gestational age (gestage):**

For gestational age, we will consider our original five categories from earlier as well as considered gestational age as a centered continuous variable and as a linear spline with a break at 38 years.

```
. gen gestage_c=gestage-38
(612 missing values generated)

. mkspline ga_sp 38 ga_sp38 = gestage, marginal
```

First the categorical version of gestational age:

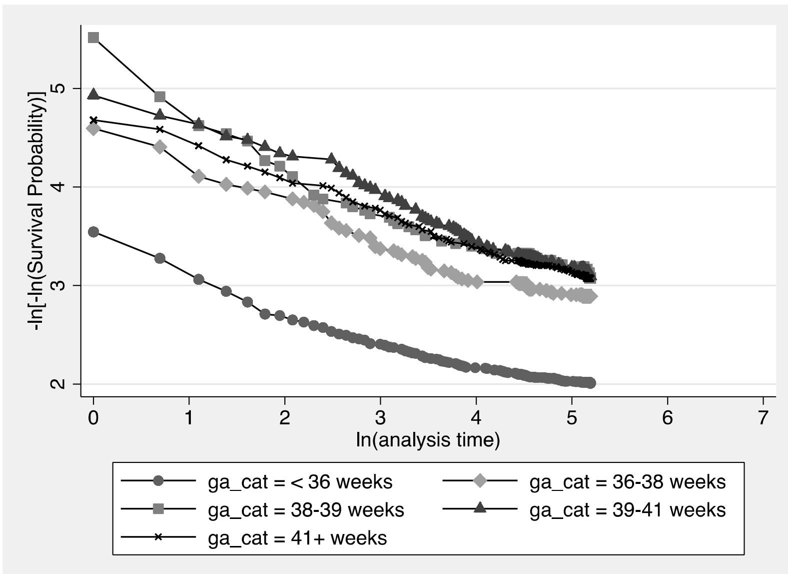
```
. stcox i.ga_cat

         failure _d:  cens == 1
   analysis time _t:  stime

Iteration 0:   log likelihood = -5439.3955
Iteration 1:   log likelihood = -5368.3146
Iteration 2:   log likelihood = -5362.5375
Iteration 3:   log likelihood = -5362.5354
Refining estimates:
Iteration 0:   log likelihood = -5362.5354

Cox regression -- Breslow method for ties

No. of subjects =        9,683                  Number of obs    =      9,683
No. of failures =          597
Time at risk    =      1542785
                                                LR chi2(4)       =     153.72
Log likelihood  =   -5362.5354                  Prob > chi2      =     0.0000

------------------------------------------------------------------------------
          _t | Haz. Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      ga_cat |
36-38 weeks  |    .413508   .0533154    -6.85   0.000     .3211698     .532394
38-39 weeks  |   .3251237   .0493452    -7.40   0.000     .2414675    .4377626
39-41 weeks  |   .3264775   .0376917    -9.70   0.000      .260365    .4093776
  41+ weeks  |   .3422158   .0403542    -9.09   0.000     .2715977    .4311952
------------------------------------------------------------------------------
```

We can also get the coefficients rather than the hazard ratios as results:

```
. stcox i.ga_cat, nohr

         failure _d:  cens == 1
   analysis time _t:  stime

Iteration 0:   log likelihood = -5439.3955
Iteration 1:   log likelihood = -5368.3146
Iteration 2:   log likelihood = -5362.5375
Iteration 3:   log likelihood = -5362.5354
Refining estimates:
Iteration 0:   log likelihood = -5362.5354

Cox regression -- Breslow method for ties

No. of subjects =        9,683                  Number of obs    =      9,683
No. of failures =          597
Time at risk    =      1542785
                                                LR chi2(4)       =     153.72
Log likelihood  =   -5362.5354                  Prob > chi2      =     0.0000

------------------------------------------------------------------------------
          _t |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      ga_cat |
36-38 weeks  |  -.8830784   .1289344    -6.85   0.000    -1.135785   -.6303715
38-39 weeks  |  -1.123549   .1517737    -7.40   0.000     -1.42102   -.8260785
39-41 weeks  |  -1.119394   .1154495    -9.70   0.000    -1.345671   -.8931173
  41+ weeks  |  -1.072314   .1179203    -9.09   0.000    -1.303433   -.8411944
------------------------------------------------------------------------------
```

We can check the proportional hazard assumption and see that the lines in the complimentary log-log plot do appear roughly parallel, although some of the hazards do cross.

```
. stphplot, by(ga_cat) ylab(2 (1) 5) xlab(0 (1) 7)

         failure _d:  cens == 1
   analysis time _t:  stime
```



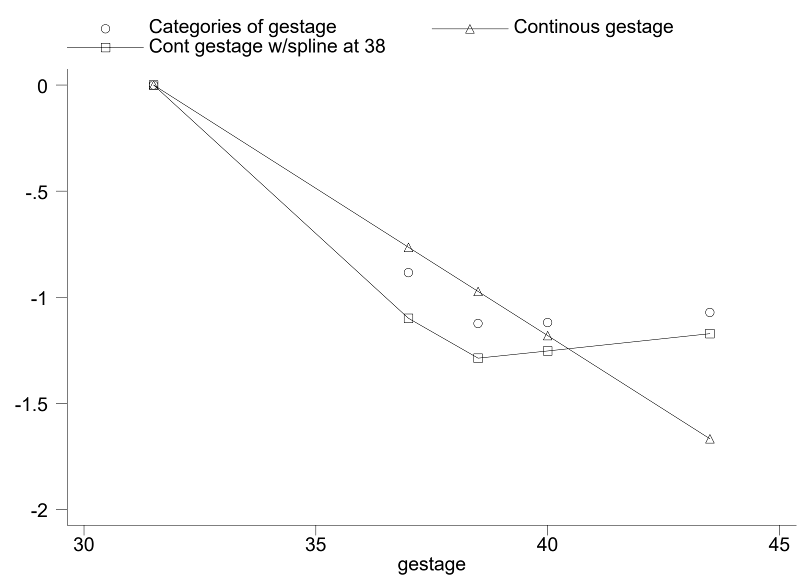
Now the centered version of gestational age:

```
. stcox gestage_c

         failure _d:  cens == 1
   analysis time _t:  stime

Iteration 0:   log likelihood = -5439.3955
Iteration 1:   log likelihood = -5359.6802
Iteration 2:   log likelihood = -5358.7331
Iteration 3:   log likelihood = -5358.7331
Refining estimates:
Iteration 0:   log likelihood = -5358.7331

Cox regression -- Breslow method for ties

No. of subjects =        9,683                   Number of obs    =       9,683
No. of failures =          597
Time at risk    =      1542785
                                                 LR chi2(1)       =      161.32
Log likelihood  =   -5358.7331                   Prob > chi2      =      0.0000

------------------------------------------------------------------------------
          _t | Haz. Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   gestage_c |   .8747498   .0089821   -13.03   0.000     .8573212    .8925327
------------------------------------------------------------------------------
```

We cannot make a complimentary log-log plot to check the proportional hazards assumption for a continuous variable, because there are no groups to compare! Finally the spline version of gestational age:

```
. stcox gestage_c ga_sp38

         failure _d:  cens == 1
   analysis time _t:  stime

Iteration 0:   log likelihood = -5439.3955
Iteration 1:   log likelihood =  -5375.191
Iteration 2:   log likelihood =  -5338.943
Iteration 3:   log likelihood = -5338.7324
Iteration 4:   log likelihood = -5338.7324
Refining estimates:
Iteration 0:   log likelihood = -5338.7324

Cox regression -- Breslow method for ties

No. of subjects =        9,683                  Number of obs    =      9,683
No. of failures =          597
Time at risk    =      1542785
                                                LR chi2(2)       =     201.33
Log likelihood  =   -5338.7324                  Prob > chi2      =     0.0000

------------------------------------------------------------------------------
          _t | Haz. Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   gestage_c |   .8185724   .0115865   -14.14   0.000     .7961754    .8415995
     ga_sp38 |   1.258339   .0440248     6.57   0.000     1.174944    1.347654
------------------------------------------------------------------------------
```

To choose between these three different models, we can consider comparing the estimated log HR of death for children in each gestational age group as compared to children in the lowest gestational age group under each of these three models, predicted at the midpoint of each gestational age category:



Both the spline and the categorical models give similar results and show the same trend. The continuous model only allows for a straight-line relationship, which seems too rigid in this case. We will include the categorical version of gestational age in our final multivariable model.

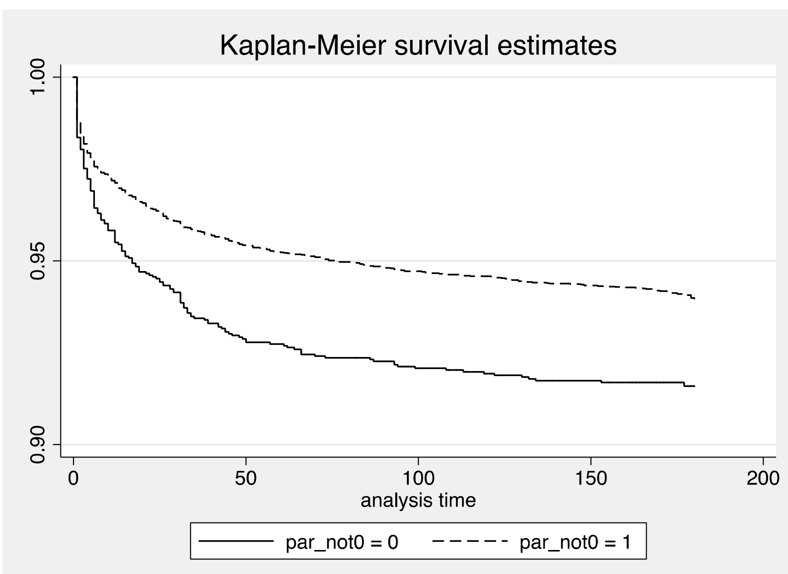
- **Parity (parity):**

For parity we will consider our original categorical version of the variable as well as a binary version which breaks into two groups: Those mothers with a previous birth (parity > 0) and those mothers where this is the first birth (parity = 0):

```
. gen par_not0=1 if parity >0 & parity < 16
(2,510 missing values generated)

. replace par_not0=0 if parity==0
(2,254 real changes made)
```

First the binary of parity. We can look at estimates of the survival curves for the two groups and see that those infants who were a first birth (parity = 0) have lower survival.

```
. sts graph, by(par_not0) ylab(.9 (.05) 1)

         failure _d:  cens == 1
   analysis time _t:  stime
```



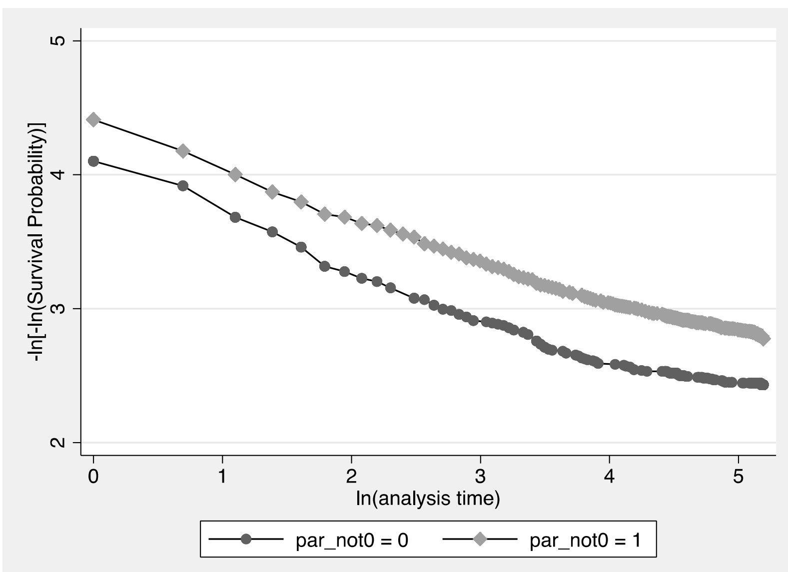
We see this in the Cox regression results as well:

```
. stcox par_not0

         failure _d:  cens == 1
   analysis time _t:  stime

Iteration 0:   log likelihood =  -5781.046
Iteration 1:   log likelihood =  -5773.349
Iteration 2:   log likelihood =  -5773.268
```

```
Iteration 3:   log likelihood =  -5773.268
Refining estimates:
Iteration 0:   log likelihood =  -5773.268

Cox regression -- Breslow method for ties

No. of subjects =       10,039                  Number of obs    =     10,039
No. of failures =          632
Time at risk    =      1600978
                                                LR chi2(1)       =      15.56
Log likelihood  =    -5773.268                  Prob > chi2      =     0.0001

------------------------------------------------------------------------------
          _t | Haz. Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    par_not0 |   .6995863   .0616632    -4.05   0.000      .588592    .8315114
------------------------------------------------------------------------------
```

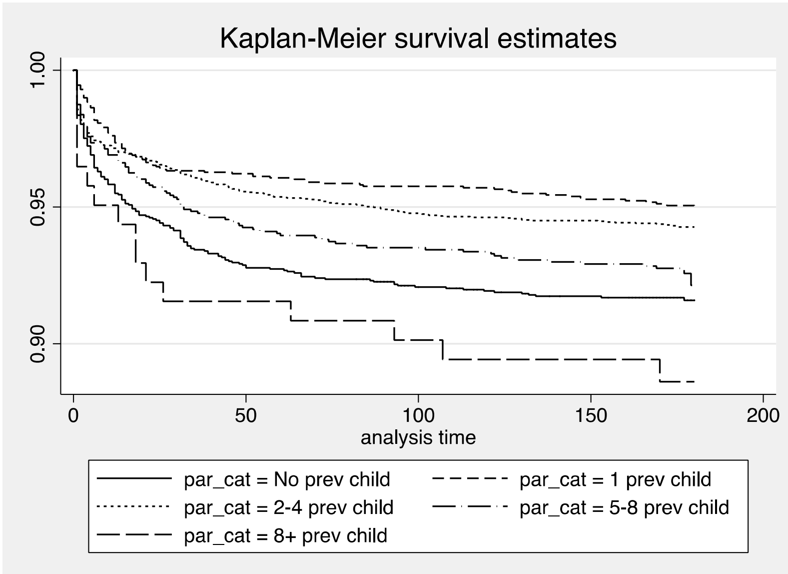
The proportional hazards assumption appears to be met!

```
. stphplot, by(par_not0) ylab(2 (1) 5) xlab(0 (1) 5)

         failure _d:  cens == 1
   analysis time _t:  stime
```



Next recall the results for the categorical version of parity:

```
. sts graph, by(par_cat) ylab(.9 (.05) 1)

         failure _d:  cens == 1
   analysis time _t:  stime
```

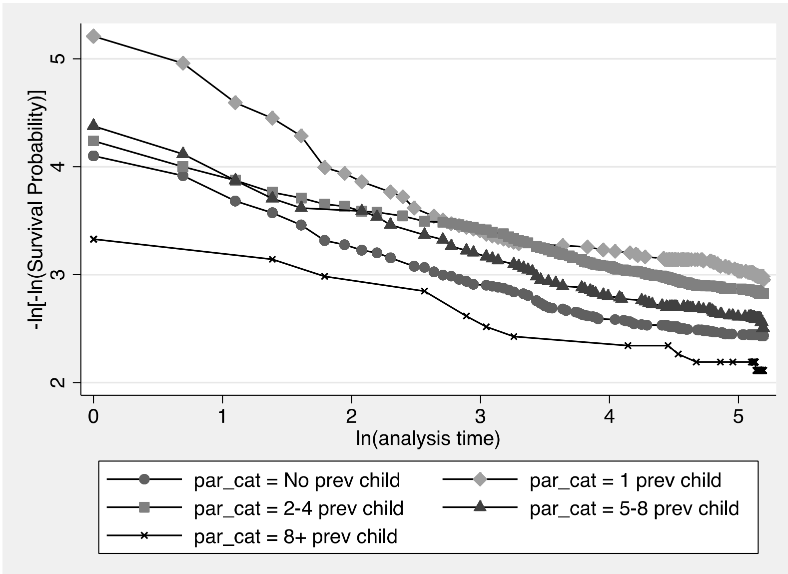


The Cox regression results are shown below:

```
. stcox i.par_cat

         failure _d:  cens == 1
   analysis time _t:  stime

Iteration 0:   log likelihood =  -5781.046
Iteration 1:   log likelihood =  -5766.178
Iteration 2:   log likelihood = -5765.8311
Iteration 3:   log likelihood = -5765.8293
Iteration 4:   log likelihood = -5765.8293
Refining estimates:
Iteration 0:   log likelihood = -5765.8293

Cox regression -- Breslow method for ties

No. of subjects =       10,039                  Number of obs    =     10,039
No. of failures =          632
Time at risk    =      1600978
                                                LR chi2(4)       =      30.43
Log likelihood  =   -5765.8293                  Prob > chi2      =     0.0000

--------------------------------------------------------------------------------
             _t | Haz. Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
----------------+---------------------------------------------------------------
        par_cat |
  1 prev child  |   .5819646   .0733035    -4.30   0.000     .4546533    .7449254
2-4 prev child  |   .6729974   .0664822    -4.01   0.000     .5545328    .8167695
5-8 prev child  |   .8852947   .1100691    -0.98   0.327     .6938372    1.129583
 8+ prev child  |   1.368832   .3570949     1.20   0.229     .8209036    2.282484
--------------------------------------------------------------------------------
```

The proportional hazards assumption again appears to be met, although the 2-4 previous children hazard crosses the other groups. We will again use the full categorical version in our multivariable Cox regression model:

```
. stphplot, by(par_cat) ylab(2 (1) 5) xlab(0 (1) 5)

         failure _d:  cens == 1
   analysis time _t:  stime
```



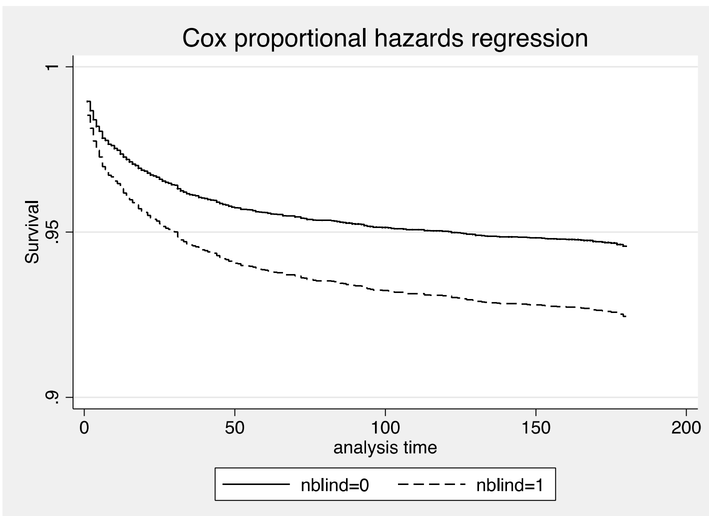
6. Finally, we build a multivariable model for survival including gestational age and parity, as well as other predictors:

```
. stcox i.ga_cat i.par_cat i.male i.nblind i.treat if male ~=9

         failure _d:  cens == 1
   analysis time _t:  stime

Iteration 0:   log likelihood = -5312.0553
Iteration 1:   log likelihood = -5222.0508
Iteration 2:   log likelihood = -5215.5804
Iteration 3:   log likelihood =  -5215.576
Iteration 4:   log likelihood =  -5215.576
Refining estimates:
Iteration 0:   log likelihood =  -5215.576

Cox regression -- Breslow method for ties

No. of subjects =        9,443                 Number of obs    =      9,443
No. of failures =          584
Time at risk    =      1523838
                                               LR chi2(12)      =     192.96
Log likelihood  =    -5215.576                 Prob > chi2      =     0.0000

--------------------------------------------------------------------------------
             _t | Haz. Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
----------------+---------------------------------------------------------------
         ga_cat |
   36-38 weeks  |   .4118807   .0537904    -6.79   0.000     .3188651    .5320296
   38-39 weeks  |   .3230919   .0495576    -7.37   0.000     .2392017    .4364033
   39-41 weeks  |   .3212073   .0377706    -9.66   0.000     .2550898    .4044619
     41+ weeks  |   .3469316   .0413817    -8.88   0.000     .2746082     .438303
                |
        par_cat |
  1 prev child  |   .5396267   .0716611    -4.65   0.000     .4159641    .7000531
2-4 prev child  |    .640755   .0654042    -4.36   0.000      .524574    .7826674
5-8 prev child  |   .7880633   .1013753    -1.85   0.064     .6124403    1.014048
 8+ prev child  |   1.179086   .3282661     0.59   0.554      .683227    2.034821
                |
           male |
```

```
         male  |   1.008702   .0836094     0.10   0.917     .8574488    1.186635
               |
        nblind |
   Night blind  |   1.407019   .1767372     2.72   0.007     1.099966    1.799783
               |
         treat |
       Placebo  |   .9556154   .0988269    -0.44   0.661     .7802871    1.170339
         Vit A  |   .9609895   .0968532    -0.39   0.693     .7887337    1.170865
--------------------------------------------------------------------------------
```

And we can look at adjusted estimates of the survival curve based on this multivariable model. For example, below we see the adjusted Kaplan-Meier survival curve by night-blindness group:

```
. stcurve, survival at1(nblind=0) at2(nblind=1) ylab(0.9 (.05) 1)
```



And the adjusted hazard curves by night-blindness group:

```
. stcurve, hazard at1(nblind=0) at2(nblind=1) yscale(log)
```