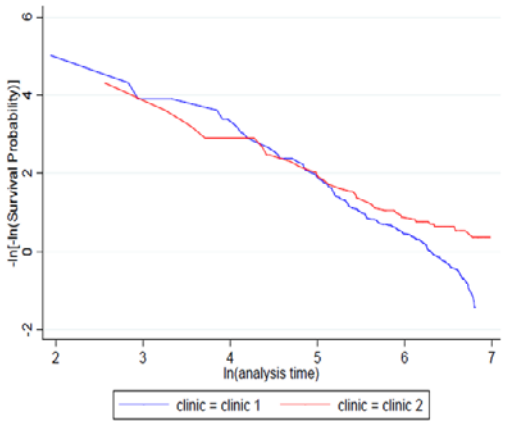
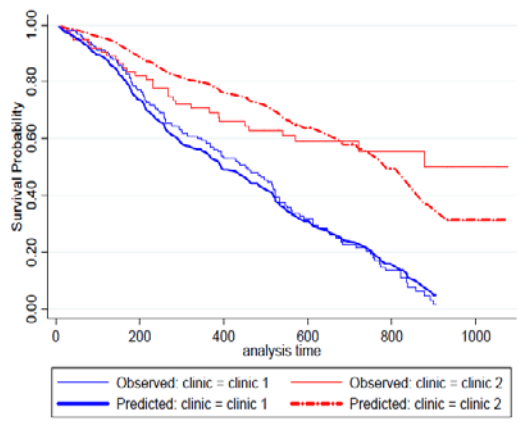
Homework 8 Solutions

1) a)
```
. stcox clinic, nohr

          failure _d:  status
    analysis time _t:  time

Cox regression -- Breslow method for ties

No. of subjects =          226                     Number of obs   =       226
No. of failures =          143
Time at risk    =        94106
                                                   LR chi2(1)      =     24.58
Log likelihood  =   -652.08842                     Prob > chi2     =    0.0000

------------------------------------------------------------------------------
          _t |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      clinic |  -.9551356   .2083378    -4.58   0.000     -1.36347   -.5468011
------------------------------------------------------------------------------

. estat phtest

      Test of proportional-hazards assumption

      Time:  Time
      ----------------------------------------------------------------
                  |                      chi2       df       Prob>chi2
      ------------+---------------------------------------------------
      global test |                     16.08        1         0.0001
```

```
. stphplot, by(clinic)
```




```
. stcoxkm, by(clinic) pred1(s(i)) pred2(s(i))
```



The PH assumption is not satisfied. We will stratify the remaining analyses by clinic.

b)Here is a table of univariate hazard ratios and 95% CI, stratifying by clinic (e.g., `. stcox female, strata(clinic)`)

Variable	Hazard ratio	Std Err.	95% CI
female	0.885	0.1732	0.604, 1.300
age	0.981	0.0198	0.943, 1.021
employ	0.646	0.1238	0.444, 0.941
benzo	0.760	0.7600	0.540, 1.069
amph	1.226	0.2790	0.785, 1.916
alc	1.235	0.2081	0.887, 1.718
dose	0.966	0.0066	0.954, 0.978

b) Here is a Stata run using all the variables

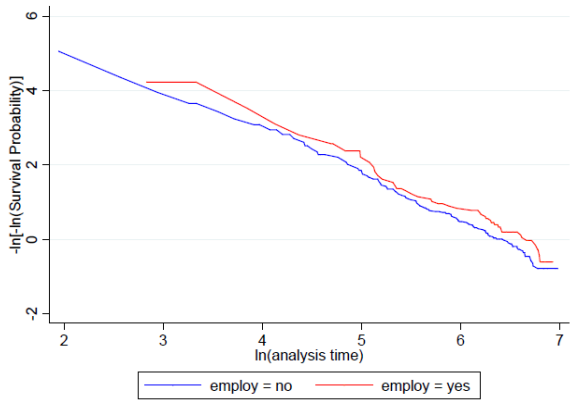
```
. stcox female age employ benzo amph alc dose, strata(clinic)

         failure _d:  status
   analysis time _t:  time

Stratified Cox regr. -- Breslow method for ties

No. of subjects =          226                     Number of obs   =       226
No. of failures =          143
Time at risk    =        94106
                                                   LR chi2(7)      =     37.92
Log likelihood  =   -555.46658                     Prob > chi2     =    0.0000

------------------------------------------------------------------------------
          _t | Haz. Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      female |   .9975521   .2097731    -0.01   0.991     .6605994    1.506375
         age |    .987356   .0205294    -0.61   0.541     .9479279    1.028424
      employ |   .6968842   .1464536    -1.72   0.086     .4616113     1.05207
       benzo |    .717695   .1309526    -1.82   0.069     .5019117    1.026248
        amph |    1.13407   .2803717     0.51   0.611     .6985547    1.841107
         alc |   1.223356   .2147707     1.15   0.251     .8671959    1.725793
        dose |   .9675486   .0063383    -5.04   0.000     .9552051    .9800515
------------------------------------------------------------------------------
                                                         Stratified by clinic
```

c) Given the investigators wished to know if any of the variables are predictive of the outcome, you might consider using a stepwise selection method. If you use a forward stepwise selection method (with entry criterion pe = 0.10) or a backward stepwise method (with entry pe = 0.10 and exit criterion pr = 0.15), the final multivariable model indicated only dose, benzodiazepine and employment status were predictive of retention. (Benzodiazepine was “borderline” significant (p-value = 0.090). It is possibly you may not have included it in your final). My model (using backward stepwise selection) was:

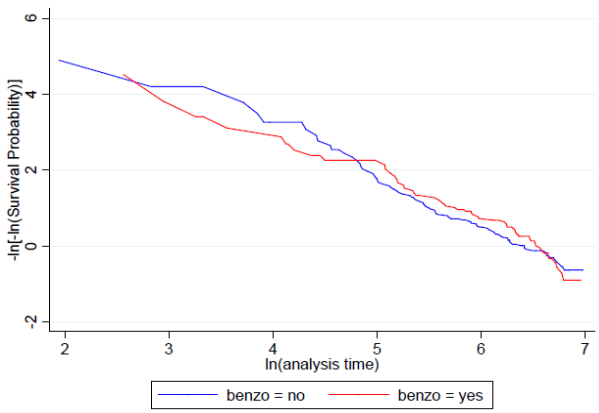
```
. sw, pe(.10) pr(.15) : stcox female age employ benzo amph alc dose, strata(clinic)
                      begin with full model
p = 0.9907 >= 0.1500  removing female
p = 0.6089 >= 0.1500  removing amph
p = 0.4872 >= 0.1500  removing age
p = 0.2527 >= 0.1500  removing alc

Stratified Cox regr. -- Breslow method for ties

No. of subjects =          226                     Number of obs   =       226
No. of failures =          143
Time at risk    =        94106
                                                   LR chi2(3)      =     35.87
Log likelihood  =   -556.48856                     Prob > chi2     =    0.0000

------------------------------------------------------------------------------
          _t | Haz. Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        dose |   .9672834   .0062659    -5.14   0.000     .9550802    .9796426
       benzo |   .7422027   .1305908    -1.69   0.090     .5257186    1.047832
      employ |   .6704283   .1308897    -2.05   0.041     .4572668     .982958
------------------------------------------------------------------------------
                                                         Stratified by clinic
```

d) Log(-log) plots for employ, benzo and (categorized) dose are

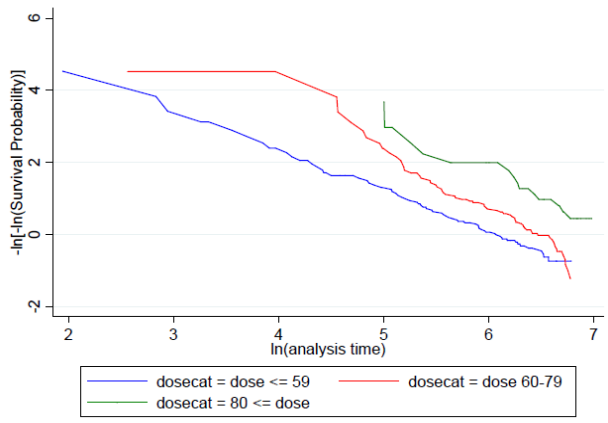
```
. stphplot, by(employ)
```




```
. stphplot, by(benzo)
```

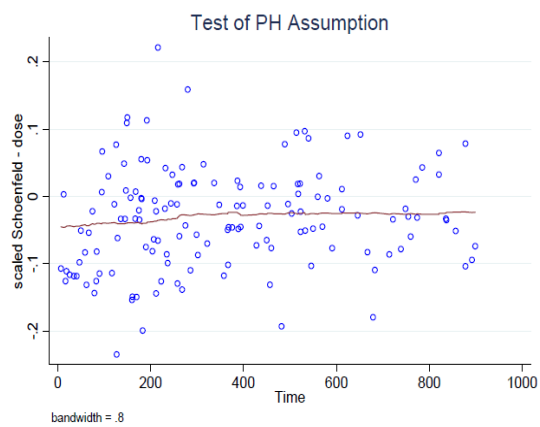


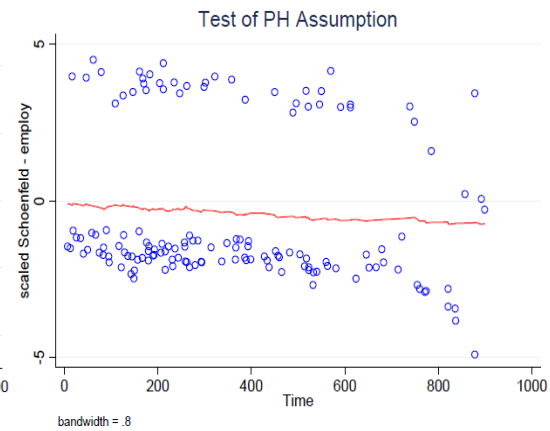

```
. stphplot, by(dosecat)  # dose categorized into a 3-level variable
```

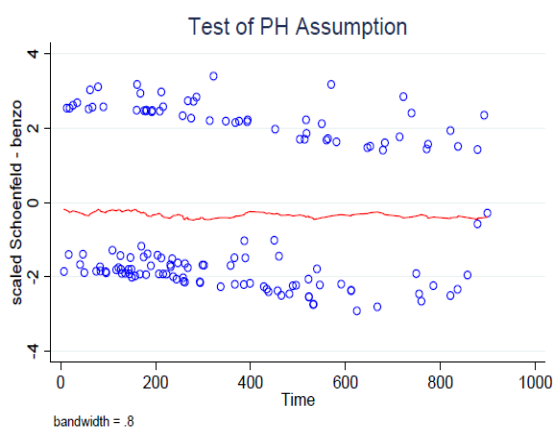


e) Checking for proportional hazards

```
. estat phtest, detail

      Test of proportional-hazards assumption

      Time:  Time
      ----------------------------------------------------------------
                  |       rho            chi2       df       Prob>chi2
      ------------+---------------------------------------------------
      dose        |      0.13613         2.45        1         0.1173
      benzo       |     -0.01067         0.02        1         0.8988
      employ      |     -0.12529         2.45        1         0.1176
      ------------+---------------------------------------------------
      global test |                      4.47        3         0.2154
      ----------------------------------------------------------------

. estat phtest, plot(dose)
. estat phtest, plot(employ)
. estat phtest, plot(benzo)
```



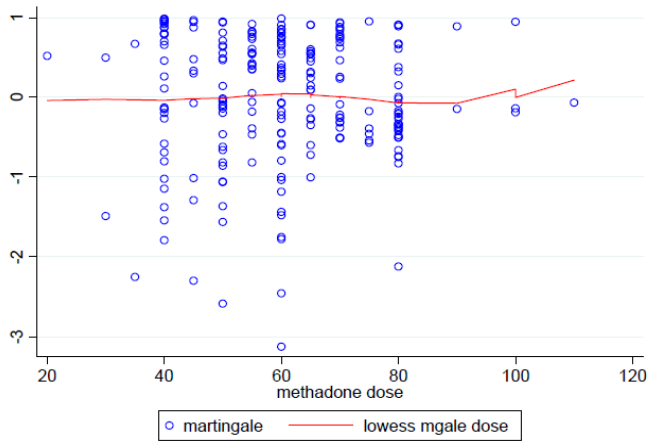




The PH assumption seems to be reasonably satisfied for all three variables.

f) To check the functional form for dose, one can plot the martingale residuals versus dose (from the model including dose) and get

```
. stcox dose benzo employ, strata(clinic) mgale(mgale)
. twoway (scatter mgale dose) (lowess mgale dose, sort)
```



The lowess line is horizontal, showing there is no meaningful trend or curvature. We can assume that the linear form in the current model is adequate.

g) The assumed model is

$h(t,X) = h_{0,clinic}(t) \exp(\beta_1 \text{ employ} + \beta_2 \text{ dose} + \beta_3 \text{ benzo})$

This implies that the relative hazards, β_1, β_2, and β_3 are the same for each clinic but the baseline hazard, h_0, varies by clinic. Other assumptions are that censoring is independent and uninformative and that the observations are independent, and we have assumed that the effect of dose is linear on the log hazard.