

# Biostatistics 209, Lab #3

## 0. Background

The purpose of this lab is to give you practice in checking proportional hazards assumption in Cox regression models.

## 1. Download Data

- Download the datasets `lab3-actg019_a.dta` and `lab3-pbc_a.dta` from the website (These are not the same datasets as you used before!).

## 2. The Data

The `lab3-actg019_a.dta` consists of two arms of ACTG 019, a study that compared zidovudine (AZT, ZDV) to placebo in a sample of asymptomatic HIV-infected subjects with CD4 count < 500 cells/mm<sup>3</sup>. The failure outcome is time from enrollment to the development of an AIDS-defining infection or death (in days).

Event Time:	<code>days</code>	Time to AIDS
Event Type:	<code>cens</code>	(0 Censoring, 1 AIDS)
Predictors:	<code>rx</code> <code>cd4</code>	Randomized Treatment ( <b>0=ZDV,1=Placebo</b> ), Baseline CD4 Count (in cells per mm <sup>3</sup> )

Declare the data to be survival time data  
`stset days, failure(cens)`

The `lab3-pbc_a.dta` contains data altered from the previous `pbc` dataset consisting of 312 subjects with primary biliary cirrhosis.

Event Time:	<code>years</code>	Time to Death
Event Type:	<code>status</code>	(0 Alive, 1 Death)
Predictors:	<code>cholest</code>	Serum Cholesterol in mg/dl

Declare the data to be survival time data  
`stset years, failure(status)`

### 3. Checking the Cox Model for ZDV treatment in lab3-actg019\_a.dta

- a. Generate the Cox-KM plot and the log-minus-log KM plot by ZDV.

```
stcoxkm, by(rx) obslopts(recast(line) connect(stairstep) lcolor(red))
obs2opts(recast(line) c(stairstep) lc(blue)) predlopts(recast(line)
c(stairstep) lc(green)) pred2opts(recast(line) c(stairstep) lc(orange))

stphplot, by(rx) nonegative nolntime plotlopts(recast(line)
connect(stairstep)) plot2opts(recast(line) c(stairstep)) ytitle(Log Minus
Log Survival) xtitle(Time) legend(order(1 "ZDV" 2 "Placebo"))
scheme(sicolor)
```

Some options used to customize the graph:

<code>recast(line)</code>	gives no symbols at the estimate points.
<code>connect(stairstep)</code>	asks Stata to connect points using a step function.
<code>lcolor(red)</code>	specifies the line color.
<code>nolntime</code>	prevents Stata from taking the log of the x-axis,
<code>nonegative</code>	specifies the log(-log) not -log(-log) plots,

- b. Does the rx HR appear proportional?
- c. Graph the log hazard ratio after fitting the Cox model and save the scaled Schoenfeld residuals. These are needed for subsequent plots and calculations. Recall, ZDV is the reference group for the variable rx.

```
stcox rx, sca(giveAname) (choose a name to be used later; no "*" with only 1 var)
estat phtest, plot(rx) (this gives a running mean smoother, rather than lowess)
```

- d. Re-graph the log hazard ratio with a lowess smoother, and save the lowess values in the variable named "`smloghr`" (obviously you can call it something else).

```
lowess giveAname days, generate(smloghr) ytitle(rx Log Hazard Ratio)
```

- e. Look at the graph and the variable `smloghr` and fill in the following table.

```
gen smhr=exp(smloghr)
sort days
list days smloghr smhr if cens==1 & (days==95|days==181|days==362|days==540)
```

	log HR	HR
95 days (3 mos)		
181 days (6 mos)		
362 days (12 mos)		
540 days (18 mos)		

- f. Run the Schoenfeld test for the proportional hazards assumption. Is there evidence against that? Does this agree with the plots?

```
estat phtest, detail
```

- g. How would you summarize the effect of ZDV on progression of HIV?

#### 4. Checking the Cox Model for Cholesterol in lab3-pbc\_a.dta

- a. Fit the Cox model for the effect of cholesterol and run the Schoenfeld test for the proportional hazards assumption. Does the evidence appear against the assumption?

```
stcox chol, sca(giveAname)           (Did you give the same name as before?)
estat phtest, detail
```

- b. Graph the log hazard ratio. What does the graph suggest? Do you have any concerns about the test?

```
lowess giveAname years, ytitle(cholesterol Log Hazard Ratio)
```

- c. Try deleting some potential influential points and then re-run the plot and test for the proportional hazards assumption. What do you conclude?

```
drop if years >12
```