EPIB 681
SOLUTIONS TO HOMEWORK #8
Question 1 – 25 points

Sharper and Fairer Comparisons: effect of sexual activity on the longevity of male fruitflies ***[For all but part h, limit analysis to fruitflies with 1 partner .. the effect is obvious in those with 8]****When we first analyzed this dataset, student PE, now on McGill faculty , argued that thorax size cannot be used as a predictor or explanatory variable since fruit flies who die young may not be fully grown, i.e., it is also an "intermediate" variable. Later, student NK (now on faculty elsewhere) had studied entomology & assured us that fruitflies do not grow longer after birth; i.e., thorax length is not time- (age)-dependent!*

a (2 points) Use a c621-type regression model (and datafile 'fruitfly') to estimate the (absolute) difference in mean longevity of sexually active flies (index category) relative to sexually inactive flies (reference category), ignoring other covariates. Is this difference (i) substantial? (ii) statistically significant at the conventional alpha=0.05 level?

```
proc ttest data=sasuser.fruitfly  ;
var lngevity;
class active;
where partners=1;
run;
```

```
                                The TTEST Procedure
                        LowerCL                            Up.CL   LowerCL                Upper CL
Variable  ACTIVE   N      Mean      Mean      Mean    Std Dev  Std Dev   Std Dev  Std Err
LNGEVITY  0       25    58.339      64.8    71.261     12.222   15.652    21.775   3.1305
LNGEVITY  1       25    50.598     56.76    62.922     11.657   14.928    20.768   2.9857
LNGEVITY  Diff (1-2)     -0.658      8.04    16.738     12.755   15.295    19.108    4.326

                                      T-Tests
             Variable    Method          Variances      DF    t Value    Pr > |t|
             LNGEVITY    Pooled          Equal          48       1.86      0.0692
             LNGEVITY    Satterthwaite   Unequal      47.9       1.86      0.0692
```

Or…

```
proc reg data=sasuser.fruitfly;
model lngevity=active;
where partners=1;
run;
```

```
                              The REG Procedure
                                  Analysis of Variance

                                          Sum of           Mean
                                       Parameter       Standard
             Variable     DF       Estimate          Error    t Value    Pr > |t|
             Intercept     1       64.80000        3.05894      21.18      <.0001
             ACTIVE        1       -8.04000        4.32600      -1.86      0.0692
```

Of course, both methods provide the same point estimates, standard errors and p-values. Both suggest that being active is not statistically significant at the 0.05 level. Even it the

difference does not reach the significance at the 5% level, this is still an important difference (big % change).

b (1 point)How different are the mean thorax lengths of the active and inactive flies? Is this difference "statistically" significant? Is statistical significance a non-issue here anyway? Explain.

```
proc ttest data=sasuser.fruitfly  ;/*b*/
var thorax;
class active;
where partners=1;
run;
```

Variable	ACTIVE	N	lower CL Mean	Mean	Upper CL Mean	Lower CL Std Dev	Std Dev	Upper CL Std Dev	Std Err
THORAX	0	25	0.7968	0.8256	0.8544	0.0546	0.0699	0.0972	0.014
THORAX	1	25	0.8085	0.8376	0.8667	0.0551	0.0706	0.0981	0.0141
THORAX	Diff (1-2)		-0.052	-0.012	0.0279	0.0586	0.0702	0.0877	0.0199

T-Tests

Variable	Method	Variances	DF	t Value	Pr > |t|
THORAX	Pooled	Equal	48	-0.60	0.5486
THORAX	Satterthwaite	Unequal	48	-0.60	0.5486

Is statistical significance a non-issue here anyway? Explain.

Yes. Confounding is not necessarily an issue of p-value. Even if the difference in thorax length is not statistically significant ($p > 0.05$), it can be clinically significant, especially if it has a large effect. Thus, we should not rely on p-values only.

c (1 point) If -- other things being equal -- flies 0.01 mm larger live on average 1 day longer, how much of a longevity "advantage" would the active flies have as a result of their larger average thorax size? On this basis, how much lower is the mean longevity of active than inactive flies if "adjusted" for the difference in thorax size?

Other things being equal, flies 0.01mm larger live on average 1 day longer and from b, sexually active flies are on average 0.01mm longer than sexually inactive flies. Therefore, sexually active fruitflies (who are on average longer) have a longer longevity by 1 day.

d (2 points)Instead of using the "out of the air" value of 1 day/ 0.01 mm, use multiple regression to simultaneously estimate the additional mean days/mm and the decrease in days associated with (due to) activity i.e., fit the model: average[longevity | thorax type] = B0 + Bthorax thorax + Bactive active

```
proc reg data=sasuser.fruitfly;
model lngevity=thorax active ;
where partners=1;
run;
```

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	1	-46.03814	20.79877	-2.21	0.0318
THORAX	1	134.25162	25.01916	5.37	<.0001
ACTIVE	1	-9.65102	*3.45570*	-2.79	0.0075

The fruitflies in the experimental group were already more advantaged at randomization since their longevity was 1.6 day longer (9.6-8.04 = 1.6) than the others.

Careful with the scales…
For each additional mm in thorax length, longevity increases by 134 days.
Same as…
Longevity is increased by 1.34 day for every 0.01mm increase in thorax length.
or
Longevity is increased by 1 day for every 1/134=0.007mm increase in thorax length.

e (1 point)Verify that if you correct the comparison in (a) using the fitted bthorax from (d) and the thorax difference in (b), you arrive at the bactive obtained in (d). Hint: cf JH notes on confounding.

Following notation in JH notes..
We have:
$Y = b0\_star + b_x\_star\ Active + b_c Thorax$
Then

b.x_star = adjusted difference
= $(Y.bar_{x=1} - Y.bar_{x=0}) - b_c(C.bar_{x=1} - C.bar_{x=0})$
= (crude differrence) – (thorax length difference)
= (from a) – (from b and d)
= (-8.04) – 134*(0.012)
= - 9.65

((134 * 0.012) corresponds to the 1.6 day correction, ie 1.5/8 = about 15% correction)

f (2 points) The p-value for the activity contrast in (d) is smaller (& the associated* CI *narrower) than the corresponding one in (a). One reason is that the larger adjusted estimate of the effect (the numerator of t-test on adjusted difference); another is the smaller SE of the estimated effect (denominator of t-test). Why is the SE of the estimated longevity difference from analysis (d) smaller?

2 reasons:

Recall:
in a: t = -8.04 / 4.32
in d: t* = -9.6 / 3.45
1- The numerator (in absolute value) has increased because of the adjustment,
2- The denominator (RMSE) has gone down because we are doing more precise (less noisy) comparisons due to the adjustment in thorax length and therefore the variance is decreased.
1 + 2 larger t, smaller p.

The longevity of the 2 groups can also be compared by survival analysis methods.
g (2 points)Use p h models to obtain "crude" & adjusted hazard ratios (again treat thorax as continuous, & use datafile 'fruitfly'). Did the adjusted coefficient move in the direction, and by the amount, you expected? Explain.

```
PROC PHREG DATA = sasuser.fruitfly;
 model lngevity*event(0) = ACTIVE  / RISKLIMITS ;
 WHERE (partners=1);
RUN;
```

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio	95% Hazard Ratio CL	
ACTIVE	1	0.46163	0.29602	2.4320	0.1189	***1.587***	0.888	2.834

```
title adjusted; /*g*/
PROC PHREG DATA = sasuser.fruitfly;
 model lngevity*event(0) = ACTIVE THORAX / RISKLIMITS ;
 WHERE (partners=1);
RUN;
```

Analysis of Maximum Likelihood Estimates

Parameter Variable	DF	Standard Estimate	Error	Chi-Square	Pr > ChiSq	Hazard Ratio	95% Hazard Ratio CL Limits	
ACTIVE	1	0.86771	0.31873	7.4115	0.0065	***2.381***	1.275	4.448
THORAX	1	-13.87820	2.77270	25.0531	<.0001	0.000	0.000	0.00

The HR increases once we adjust for thorax length, i.e. the death rate in sexually active fruit flies is increased as compared to the death rate in inactive fruitflies. In previous section, we found that longevity in sexually active fruitflies was even shorter after adjustment. Therefore the 2 results are concordant (effect is magnified).

h (2 points) Repeat question g, but for flies with 8 partners: Do the adjustments go in the direction you expected? Explain.

```
proc ttest data=sasuser.fruitfly  ;/*a*/
var longevity thorax;
class active;
where partners=8;
run;
```

The TTEST Procedure

Statistics

Variable	ACTIVE	N	Lower CL Mean	Mean	Upper CL Mean	Lower CL Std Dev	Std Dev	Upper CL Std Dev	Std Err
LNGEVITY	0	25	57.358	63.36	69.362	11.353	14.54	20.227	2.908
LNGEVITY	1	25	33.725	38.72	43.715	9.4496	12.102	16.836	2.4204
LNGEVITY	Diff (1-2)		17.033	24.64	32.247	11.155	13.377	16.711	3.7835
THORAX	0	25	0.7719	0.8056	0.8393	0.0637	0.0816	0.1135	0.0163
THORAX	1	25	0.7677	0.8	0.8323	0.0612	0.0783	0.1089	0.0157
THORAX	Diff (1-2)		-0.04	0.0056	0.0511	0.0667	0.0799	0.0999	0.0226

T-Tests

Variable	Method	Variances	DF	t Value	Pr > |t|
LNGEVITY	Pooled	Equal	48	6.51	<.0001
LNGEVITY	Satterthwaite	Unequal	46.5	6.51	<.0001
THORAX	Pooled	Equal	48	0.25	0.8055
THORAX	Satterthwaite	Unequal	47.9	0.25	0.8055

```
title crude; /*h*/
PROC PHREG DATA = sasuser.fruitfly;
 model lngevity*event(0) = ACTIVE   / RISKLIMITS ;
 WHERE (partners=8);
RUN;
```

Analysis of Maximum Likelihood Estimates

Parameter variable	DF	Standard Estimate	Error	Chi-Square	Pr > ChiSq	Hazard Ratio	95% HR ConfLimits	
ACTIVE	1	2.08297	0.42982	23.4853	<.0001	***8.028***	3.457	18.642

```
title adjusted; /*h*/
PROC PHREG DATA = sasuser.fruitfly;
 model lngevity*event(0) = ACTIVE THORAX / RISKLIMITS ;
 WHERE (partners=8);
RUN;
```

Analysis of Maximum Likelihood Estimates

Parameter Variable	DF	Standard Estimate	Error	Chi-Square	Pr > ChiSq	Hazard Ratio	95% CL	
ACTIVE	1	3.22402	0.51920	38.5591	<.0001	***25.129***	9.083	69.521
THORAX	1	-18.30515	3.08731	35.1550	<.0001	0.000	0.000	0.000

Do the adjustments go in the direction you expected? Explain.

No since experimental are shorter to begin with. HR is increased after adjustment, even though sexually active have a shorter longevity to start with.

This is a Particular phenomenon in logistic and other non linear models, ie, confounding on the additive scale (MLR doe not imply confounding on the multiplicative scale

(logistic, PH..), and vice-versa. Ie, cannot predict change in estimates in PH from results of linear regression analyses.

i (1 point)Use logs to write the p h models so that the right hand side has the same additive forms as in (a) and (d)

$h(t|\ \text{Active, thorax}) = h_0(t)\ \exp(\beta_1 \text{Active} + \beta_2\ \text{Thorax})$

$\log(h(t|\ \text{Active, thorax})) = \log[h_0(t)] + \beta_1 \text{Active} + \beta_2\ \text{Thorax}$

j (2 points) Some flies began adult life on Mondays, some on Tuesdays etc. The research assistant entered data for each fly each Saturday, making a separate (partial) record for each week, or part thereof: for example, a fly who began as an adult on Tuesday, and ultimately lived 62 days, has 10 records -- each of the first 9, namely those where (t0,t1) = (0,4), (4,11) to.(53,60) is accompanied by a "censored" indicator, and a 10th record, with (t0,t1) =(60,62) is accompanied by a "complete" indicator (Run SAS/Stata program and inspect the "byweek " datafile created). Repeat the crude & survival comparisons with these split observations. Explain why the likelihoods, beta_hats, SE's etc. are identical to those in (g). A diagram, with timelines & risksets, may help.

```
title 'weekly records';
PROC PHREG DATA = sasuser.byweek;/*j*/
 model (t_0,t_1)*complete(0) = ACTIVE   / RISKLIMITS;
 WHERE (partners=1);
RUN;
```

```
                                Model Information
                          Data Set                  SASUSER.BYWEEK
                          Dependent Variable        t_0
                          Dependent Variable        t_1
                          Censoring Variable        complete
                          Censoring Value(s)        0
                          Ties Handling             BRESLOW

                       Summary of the Number of Event and Censored Values
                         Total       Event    Censored   % Censored
                           484          50         434       89.67

                         Analysis of Maximum Likelihood Estimates
Parameter    Standard                                          Hazard   95% Hazard Ratio
Variable  DF    Estimate       Error  Chi-Square  Pr > ChiSq     Ratio   Conf Limits
ACTIVE     1    0.46163     0.29602      2.4320      0.1189     1.587     0.888     2.834
```

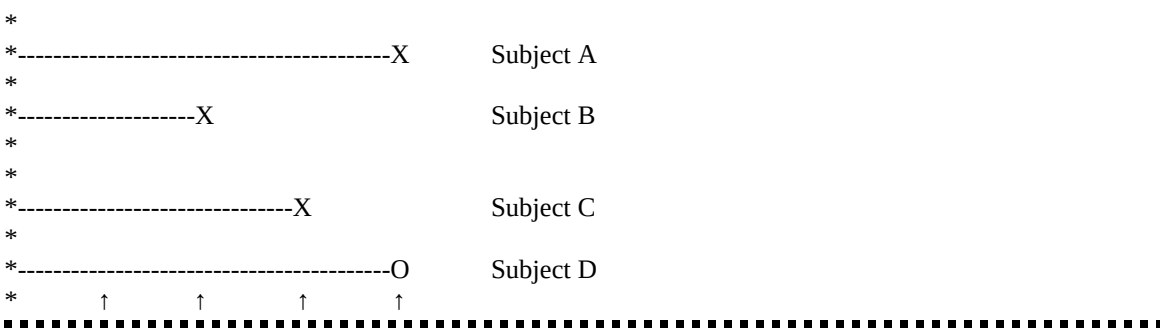
```
title 'weekly records';
PROC PHREG DATA = sasuser.byweek;/*j*/
 model (t_0,t_1)*comlete(0) = ACTIVE THORAX / RISKLIMITS;
 WHERE (partners=1);
RUN;
```

```
            Analysis of Maximum Likelihood Estimates

Parameter       Standard                                          Hazard   95% Hazard Ratio
Variable  DF    Estimate       Error  Chi-Square  Pr > ChiSq     Ratio   ConfLimits
ACTIVE     1    0.86771     0.31873      7.4115      0.0065     2.381     1.275     4.448
THORAX     1   -13.87820     2.77270     25.0531      <.0001     0.000     0.000     0.000
```

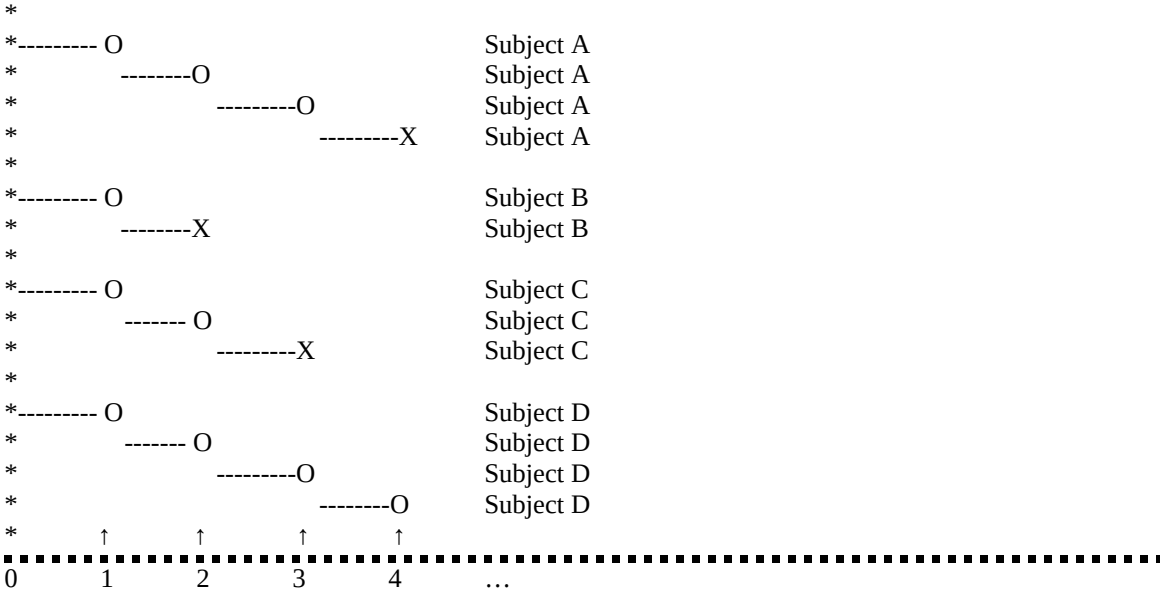
The proportional hazard likelihood is computed at each risk set. I.e., in each risk set, the numerator includes all the fruitflies who died at that time, and the denominator includes all the fruitflies who are in the risk set at that time (those still alive at that time + those who dies at that time). Therefore, when we split the time, we are not creating new events. The likelihood will be the same for both methods given that the same cases and the same non-cases will be in the risk-sets.

Ie: (X: died, O: alive)



time 1: no risk set
In the risk set at time 2: A, B, C, D (1 case, 3 non-cases)
In the risk set at time 3: A, B, D (1 case, 2 non-cases)
In the risk set at time 4: A, D (1 case, 1 non-case)

SAME AS:



time 1: no risk set
In the risk set at time 2: A, B, C, D (1 case, 3 non-cases)
In the risk set at time 3: A, B, D (1 case, 2 non-cases)
In the risk set at time 4: A, D (1 case, 1 non-case)

k <u>(2 points)</u> On the weekend, the RA's boss, not knowing how to get SAS/Stata to put the 1154 split records back together, decided that one should only use the 50 split records that terminated in a death (see near the end of the program file). Run the boss's analysis, and explain why it gives a very different (wrong?) answer.

```
title weekly records, but only the week when the event occurred;
title2 (The BOSSs analysis);
PROC PHREG DATA = sasuser.byweek;
 model (t_0,t_1)*complete(0) = ACTIVE THORAX / RISKLIMITS;
 WHERE (partners=1 and complete=1);
RUN;
```

```
                    Summary of the Number of Event and Censored Values
                          Total       Event    Censored    %Censored
                             50          50           0        0.00
                        Analysis of Maximum Likelihood Estimates
Parameter    Standard                                          Hazard   95% Hazard Ratio
Variable  DF    Estimate       Error  Chi-Square  Pr > ChiSq     Ratio   ConfLimits
ACTIVE    1     0.20455     0.39480      0.2684      0.6044     1.227     0.566     2.660
THORAX    1    -4.11288     3.20248      1.6494      0.1990     0.016     0.000     8.705
```

If we look back at the LKD, in each risk set we now have the same numerator (those who dies at the time j) but the denominator does not include all the fruitflies anymore. Ie, we are underestimating the denominator of the LKD and for that reason we get completely different (wrong) results. We are looking at ratio of death rates, therefore, at longevity. In this case, we should take into account the complete longevity, i.e., all the time between birth to death.

If we only take last week of data, we are basically using as “time 0“ the first day of the week that preceded death. This is meaningless….

l <u>(2 points)</u> Rather than use thorax size as a term in (and a coefficient to be estimated from) the regression, use the variable thorax_Q (quintiles 1 to 5) as a stratum variable and re-estimate the HR for the active relative. to the inactive group. Why is the likelihood much larger (the logL less negative) in the stratified analysis? Hint: examine sizes of the risksets, and likelihood contributions, in Figures 1 and 3 of JH's draft article on Survival analysis; risk sets; case control studies: part II.

```
title Thorax STRATA rather than as a regressor variable;
PROC PHREG DATA = sasuser.byweek;
 model (t_0,t_1)*complete(0) = ACTIVE / RISKLIMITS ;
 STRATA thorax_q;
 WHERE (partners=1);
RUN;
```

Model Fit Statistics

Criterion	Without Covariates	With Covariates
-2 LOG L	155.492	149.723
AIC	155.492	151.723
SBC	155.492	153.635

Summary of the Number of Event and Censored Values

Stratum	Thorax_Q	Total	Event	Censored	Percent Censored
1	1	75	10	65	86.67
2	2	70	8	62	88.57
3	3	125	12	113	90.40
4	4	109	10	99	90.83
5	5	105	10	95	90.48
Total		484	50	434	89.67

Analysis of Maximum Likelihood Estimates

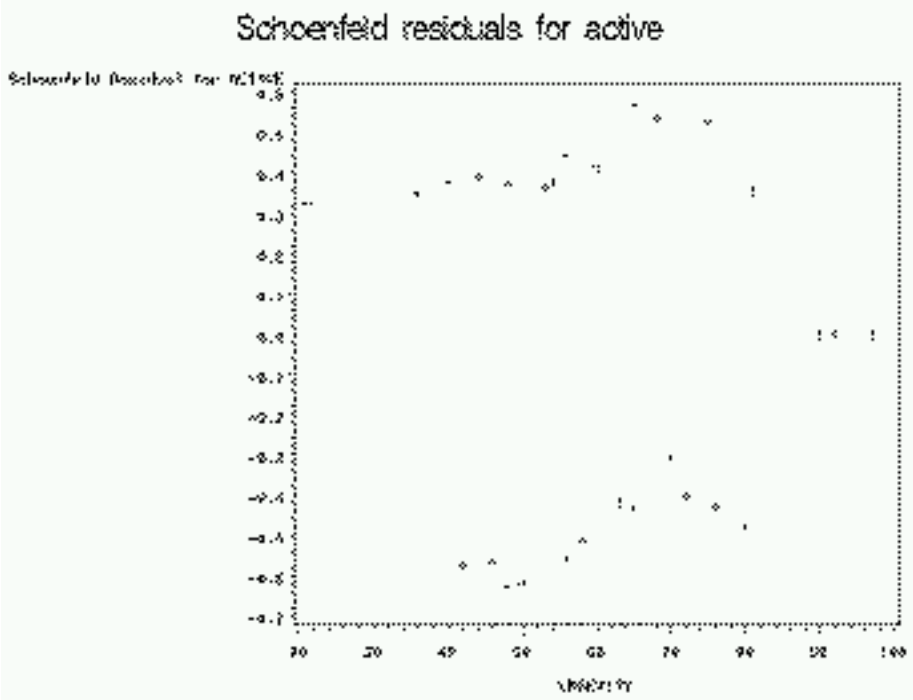
Variable	DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio	95% Hazard Ratio ConfLimits	
ACTIVE	1	0.79908	0.33657	5.6369	0.0176	2.223	1.150	4.301

From JH draft article *'The different log-likelihood scale, compared with Figure 2, stems from the fact that each riskset is smaller, so that the associated probability is larger, and the log-probability is less negative. For this reason, the log-likelihood based on these stratified series cannot be compared with the log-likelihood from the 2-parameter model.'*

m <u>(3 points)</u>Use the Schoenfeld residuals, & log[-Log[S]] plots, to visually assess if the p h assumption is reasonable in his dataset. (consult onlinedoc or other documentation)

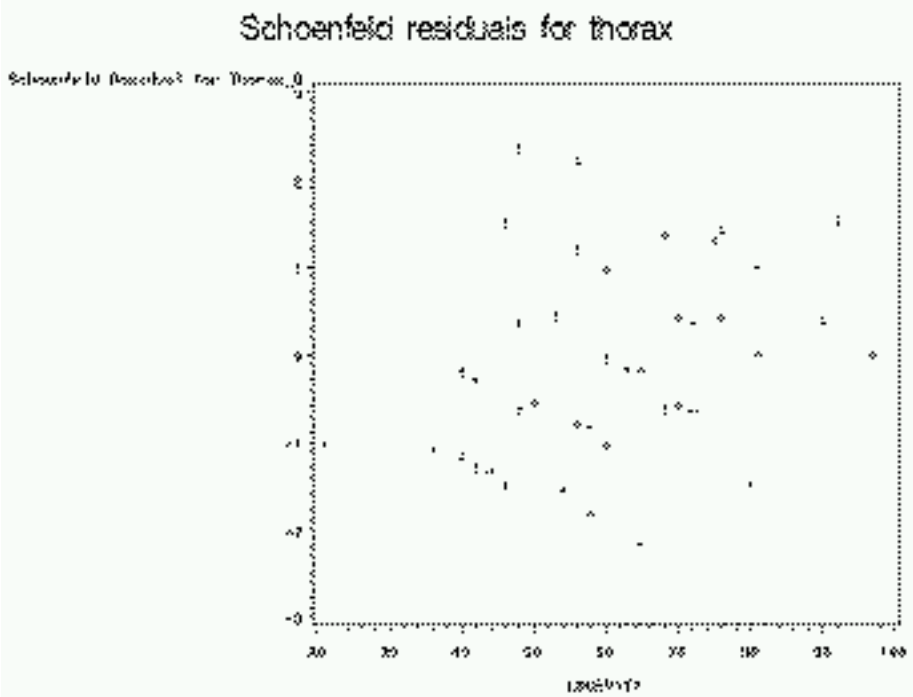
Schoenfeld residuals

These are obtained using **`proc`** `phreg` for ***<u>each</u>*** covariate. They have to be saved and then plotted against time.



?? pattern of time dependency.

```
PROC PHREG DATA = sasuser.fruitfly  ;
 model lngevity*event(0) = thorax_q ACTIVE / RISKLIMITS ;
 WHERE (partners=1);
 output out=sch_thorax ressch=schoenfeld ;
RUN;
proc gplot data=sch_thorax;
title 'Schoenfeld residuals for thorax';
plot schoenfeld * lngevity;
run;
quit;
```

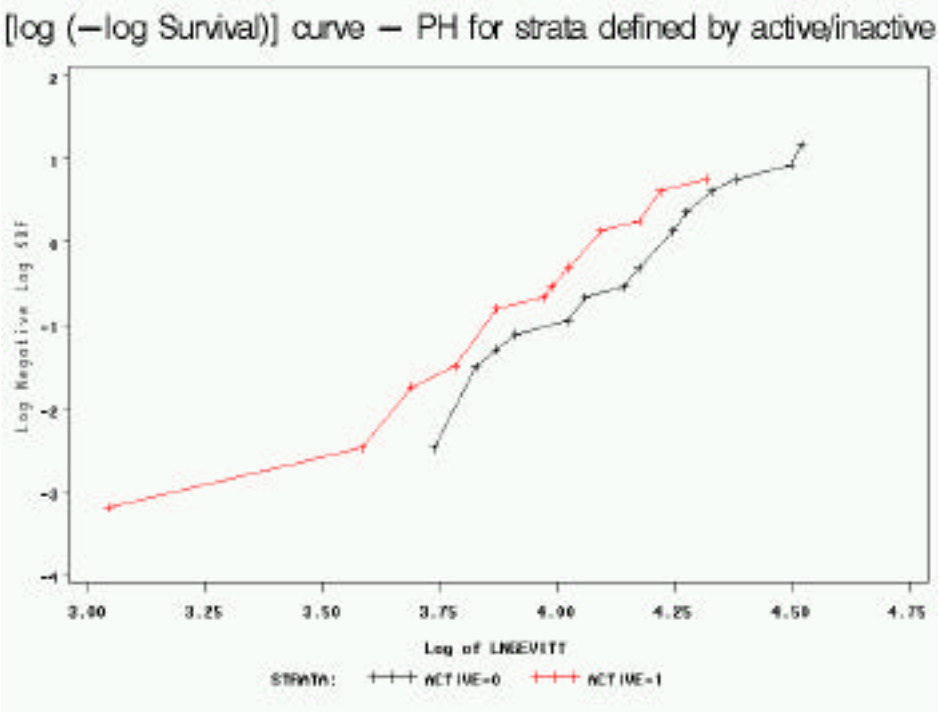


no pattern of time dependency.

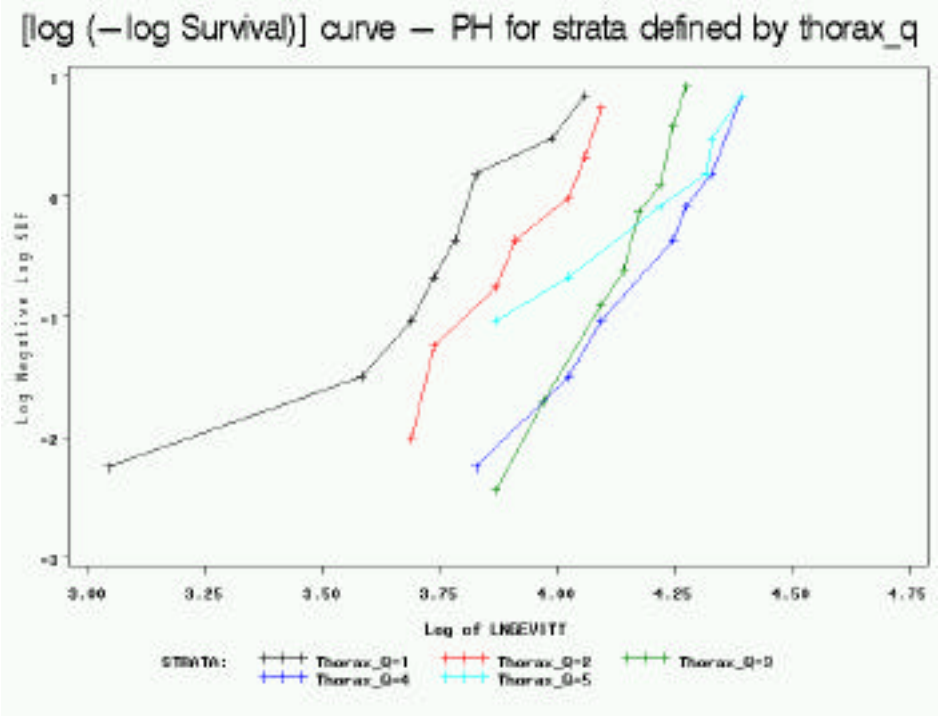
log[-Log[S]] plots

```
proc lifetest data=sasuser.fruitfly plots=(lls);
time lngevity*event(0);
STRATA active;
WHERE (partners=1);
run;
proc sort data=mydata;by thorax_q;

proc lifetest data=sasuser.fruitfly plots=(lls);
time lngevity*event(0);
STRATA thorax_q;
WHERE (partners=1);
run;
```



Assumption of PH looks reasonable for this variable (lines approximately parallel).



Assumption of PH looks reasonable for this variable (lines approximately parallel, we should not be too picky, sample sizes are small…)

<u>Shoenfeld in stata:</u>

from : http://www.ats.ucla.edu/stat/sas/faq/test_proportionality.htm

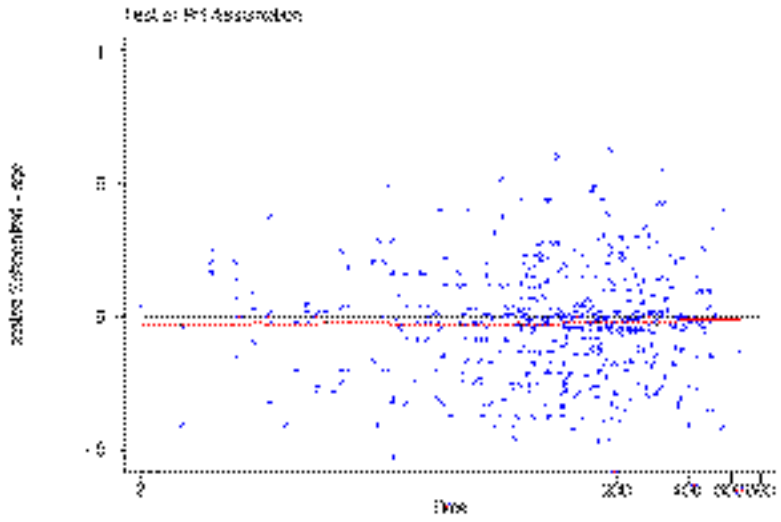
Tests and graphs based on the Schoenfeld residuals

Testing the time dependent covariates is equivalent to testing for a non-zero slope in a generalized linear regression of the scaled Schoenfeld residuals on functions of time. A non-zero slope is an indication of a violation of the proportional hazard assumption. As with any regression it is highly recommended that you look at the graph of the regression in addition to performing the tests of non-zero slopes. There are certain types on non-proportionality that will not be detected by the tests of non-zero slopes alone but that might become obvious when looking at the graphs of the residuals such as nonlinear relationship (i.e., a quadratic fit) between the residuals and the function of time or undue influence of outliers. For more details please refer to Modeling Survival Data by Therneau and Grambsch p. 127-142.

Stata The tests of the non-zero slope developed by Therneau and Grambsch for SPLUS have been implemented in STATA in the **stphtest** command. The algorithms that STATA uses are slightly different from the algorithms used by SPLUS and therefore the results from the two programs might differ slightly. The **stphtest** with the **detail** option will perform the tests of each predictor as well as a global test. There are different functions of time available including the identity function, the log of survival time and the rank of the survival times. The **stphtest** command with the **plot** option will provide the graphs with a lowess curve. The usual graphing options can be used to include a horizontal reference line at y=0. Unlike the graphs created in SPLUS the graphs in Stata do not include 95% confidence intervals for the lowess curves which makes it more difficult to assess how much the curves may deviate from the y=0 line.

stcox age race treat site agesite, nolog noshow schoenfeld(sch*) scaledsch(sca*)
stphtest, log detail
stphtest, log plot(age) yline(0)
stphtest, log plot(treat) yline(0)

Test of proportional hazards assumption



n (1 point) What role should the variable SLEEP have in this analysis? Based on the data, does it seem to be influential/relevant?

Sleep corresponds to % of time sleeping. We could expect that the flies that sleep more are less active… However it should not affect thorax length nor longevity…

o (2 points) Compare beta_hats in (g) with those from "ran_out" dataset. Explain why the larger SE's, and by how much.

Study was ended earlier at day 84 and therefore the period of observation is shorter (like in an RCTm, where sometimes investigators decide that study will end at the same date for everybody). The parameter estimates are not biased, however, because the total FU time is reduced, the estimates are less precise (the standard errors are bigger, here we have 0.48 and before we had 0.31)

```
title funding ran out;
PROC PHREG DATA = sasuser.ran_out;
 model t_fu*final(0) = ACTIVE THORAX / RISKLIMITS;
 WHERE (partners=1);
RUN;
```

Model Fit Statistics

Criterion	Without Covariates	With Covariates
-2 LOG L	148.969	123.139
AIC	148.969	127.139
SBC	148.969	129.577

Analysis of Maximum Likelihood Estimates

Parameter Variable	DF	Standard Estimate	Error	Chi-Square	Pr > ChiSq	Hazard Ratio	95% Hazard Ratio ConfLimits	
ACTIVE	1	1.45806	*0.48381*	9.0826	0.0026	4.298	1.665	11.093
THORAX	1	-18.15518	3.87233	21.9814	<.0001	0.000	0.000	0.000