

Biostatistics 209

Homework #2

Due April 22, 2014: *hardcopy to Olivia De Leon by start of class*

For Question 3, use the dataset `pbcc.dta` available on the website.

1. [4 points] The follow Stata output gives the hazard ratio of a **1-year increase** in age (variable named `agedx`).

_t	Haz. Ratio	Std. Err.	z	P> z	[95% Conf. Interval]
agedx	.9144948	.0264917	-3.09	0.002	.8640186 .9679198

Based on the above, with a **1-year decrease** in age, what are the corresponding

- (a) [1 point] Hazard ratio,
 - (b) [2 points] Wald test statistic & p-value, and
 - (c) [1 point] 95% CI?
2. [4 points] The variable `agedx` was grouped as follows to create the variable “ages”

agedx: 0-2: ages =1

agedx: 2-6: ages =2

agedx: 6-12: ages =3

agedx: > 12: ages =4

This leads the following output from a Cox model

```

No. of subjects =          103          Number of obs   =          103
No. of failures =           74
Time at risk    =  1413.050005
Log likelihood   = -266.66906
LR chi2(3)       =          28.55
Prob > chi2      =          0.0000

```

_t	Haz. Ratio	Std. Err.	z	P> z	[95% Conf. Interval]
agedx					
2	.2684887	.1685833	-2.09	0.036	.0784258 .9191639
3	.0989741	.0632982	-3.62	0.000	.028258 .346658
4	.0566843	.0391472	-4.16	0.000	.0146422 .2194414

- (a) [1 point] What is the hazard ratio for someone aged 0-2 (ages =1) compared to someone aged >12 (ages =4)? *i.e., now considering (ages =4) as the reference group.*

- (b) [2 points] What are the Wald test statistic, p-value and 95% CI associated with the hazard ratio in (a)?
- (c) [1 point] What is the hazard ratio for someone aged 2-6 (ages =2) compared to someone aged >12 (ages =4)? *i.e., considering (ages =4) as the reference group here.* (Hint: look at the definition of hazard ratio for 2 versus 1 and 4 versus 1)
3. [5 points] For the variable age in the pbc dataset, check the proportional hazards assumption.
- (a) [1 point] Using `stphplot` (log-log plot) or `stcoxkm`. *Note, for this question, you will need to create a grouped version of age. Use the Stata commands:*
- ```
egen Gage=cut(age), at(25,40,50,60,80) label
label var Gage "Age Group"
tab status Gage
```
- (b) [1 point] Using Smoothed hazard ratios for the variable age.
- (c) [1 point] Perform the Schoenfeld test for the variable age.
- (d) [2 points] Based on (a)-(c), what do you conclude? Are hazards for the variable age proportional? How would you report the effect (in terms of HR and 95%CI) of age in a paper?