Biostatistics 209 Homework #2
Due April 21, 2009 *by hardcopy in class*

1. 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 **5-year** increase in age, what are the corresponding

   (a) Hazard ratio,

   (b) Wald test & p-value, and

   (c) 95% CI?

2. 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
                                                    LR chi2(3)       =     28.55
Log likelihood  =   -266.66906                      Prob > chi2      =    0.0000

------------------------------------------------------------------------------
          _t | Haz. Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    _Iages_2 |   .2684887   .1685833    -2.09   0.036     .0784258    .9191639
    _Iages_3 |   .0989741   .0632982    -3.62   0.000      .028258     .346658
    _Iages_4 |   .0566843   .0391472    -4.16   0.000     .0146422    .2194414
------------------------------------------------------------------------------
```

   (a) What is the hazard ratio for someone aged 0-2 (`ages` =1) compared to someone aged 6-12 (`ages` =3)? *Note, consider* (`ages` =3) *as the reference group here.*

   (b) Give the Wald test, p-value and 95% CI associated with the hazard ratio in (a)?

   (c) What is the hazard ratio for someone aged 2-6 (`ages` =2)? compared with someone aged 6-12 (`ages` =3)? *Note, again consider* (`ages` =3) *as the reference group here.*

For Question 3, use the dataset pbc.dta available on the course website.

3. For the variable `age` in the pbc dataset, check the proportional hazards assumption.

   (a) Using log-minus-log plots. *Note, you will need to create a grouped version of* `age`. *Use the* `Stata` *commands* `xtile` *or* `cut`.

   (b) Smoothed hazard ratios.

   (c) The Schoenfeld test.

   (d) Based on (a)-(c), what do you conclude? Are hazards for `age` proportional? How would you report the effect of `age` in a paper?