Interactions - Exercise 111(j and k)

Elisavet Syriopoulou (elisavet.syriopoulou@ki.se)

January 18, 2022

Here is the output from the fitted model, with the model formulation written below. Note that the output doesn't give us the β-coefficients but $\exp(\beta)$, since it says Haz. Ratio at the top of the table. This is the default for the command streg, but other commands could have the default of showing the β-coefficients, so make sure to check.

```
. streg i.fu i.agegrp i.year8594##i.sex, dist(exp)

------------------------------------------------------------------------------------------
                     _t | Haz. Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
------------------------+-----------------------------------------------------------------
                     fu |
                     1  |   3.554795   .4831838     9.33   0.000     2.723425    4.639955
                     2  |   3.693547   .5099324     9.46   0.000     2.817906    4.841287
                     3  |   2.932013   .4288725     7.35   0.000     2.201195    3.905468
                     4  |   2.447604   .3808316     5.75   0.000     1.804262    3.320341
                     5  |    2.25602   .3692772     4.97   0.000     1.636868    3.109367
                     6  |   1.797325   .3227558     3.26   0.001     1.264071    2.555534
                     7  |   1.288401    .267454     1.22   0.222     .8577355    1.935301
                     8  |   1.439152   .3023187     1.73   0.083     .9534478    2.172282
                     9  |   .7958958    .221615    -0.82   0.412     .4611492    1.373634
                        |
                 agegrp |
                 45-59  |   1.326709   .1249663     3.00   0.003     1.103059    1.595705
                 60-74  |   1.861131   .1691561     6.83   0.000     1.557443    2.224035
                   75+  |   3.399539   .3550374    11.72   0.000     2.770277    4.171737
                        |
               year8594 |
         Diagnosed 85-94  |   .7414351   .0655414    -3.38   0.001     .6234888    .8816936
                        |
                    sex |
                Female  |   .6031338   .0531555    -5.74   0.000     .5074526     .716856
                        |
           year8594#sex |
Diagnosed 85-94#Female  |   .9437245   .1232639    -0.44   0.657     .7305772    1.219058
                        |
                  _cons |   .0125379     .00183   -30.00   0.000     .0094185    .0166904
------------------------------------------------------------------------------------------
```

$$\ln(\lambda) = \beta_0+\beta_1\text{fu}_{1-2}+\beta_2\text{fu}_{2-3}+\beta_3\text{fu}_{3-4}+\beta_4\text{fu}_{4-5}+\beta_5\text{fu}_{5-6}+\beta_6\text{fu}_{6-7}+\beta_7\text{fu}_{7-8}+\beta_8\text{fu}_{8-9}+\beta_9\text{fu}_{9-10}$$
$$+\beta_{10}\text{age}_{45-59}+\beta_{11}\text{age}_{60-74}+\beta_{12}\text{age}_{75+}+\beta_{13}\text{year8594}+\beta_{14}\text{female}+\beta_{15}(\text{year8594}\times\text{female})$$

On the rate scale this can be written as:

$$\begin{aligned}\lambda = \exp(&\beta_0+\beta_1\text{fu}_{1-2}+\beta_2\text{fu}_{2-3}+\beta_3\text{fu}_{3-4}+\beta_4\text{fu}_{4-5}+\beta_5\text{fu}_{5-6}+\beta_6\text{fu}_{6-7}+\beta_7\text{fu}_{7-8}+\beta_8\text{fu}_{8-9}+\beta_9\text{fu}_{9-10}\\ &+\beta_{10}\text{age}_{45-59}+\beta_{11}\text{age}_{60-74}+\beta_{12}\text{age}_{75+}+\beta_{13}\text{year8594}+\beta_{14}\text{female}+\beta_{15}(\text{year8594}\times\text{female}))\end{aligned}$$

We want to estimate the mortality rate ratio of sex for patients for each calendar period.

Let's look at the first year of follow-up (reference group for fu) and age-group 0-44 (reference group for age).
First, for those diagnosed during 1975−84 (year8594=0).
The rate for females (female=1):

$$\exp(\beta_0 + \beta_{14})$$

The rate for males (female=0):

$$\exp(\beta_0)$$

The mortality rate ratio of females compared to males for the earlier calendar period will be:

$$\frac{\exp(\beta_0 + \beta_{14})}{\exp(\beta_0)} = \frac{\exp(\beta_0)\exp(\beta_{14})}{\exp(\beta_0)} = \exp(\beta_{14}) = 0.6031338$$

Similarly, we can estimate the mortality rate ratio of sex for patients diagnosed during 1985−94 (year8594=1).
The rate for females (female=1):

$$\exp(\beta_0 + \beta_{13} + \beta_{14} + \beta_{15})$$

The rate for males (female=0):

$$\exp(\beta_0 + \beta_{13})$$

The mortality rate ratio of females compared to males for the later calendar period will be:

$$\begin{aligned}\frac{\exp(\beta_0 + \beta_{13} + \beta_{14} + \beta_{15})}{\exp(\beta_0 + \beta_{13})} = \frac{\exp(\beta_0)\exp(\beta_{13})\exp(\beta_{14})\exp(\beta_{15})}{\exp(\beta_0)\exp(\beta_{13})} &= \exp(\beta_{14})\exp(\beta_{15})\\ &= 0.6031338 \times 0.9437245 = 0.569192\end{aligned}$$

So we showed the hazard ratio comparing males to females in both calendar period, but only within the first year since diagnosis and the youngest age group. We want to know what the hazard ratio is for females compared to males, in both calendar periods, within all follow-up years and age groups. Let's now look at the second year of follow-up ($\text{fu}_{1-2} = 1$) and age-group 45-59 ($\text{age}_{45-59} = 1$). First for patients diagnosed during 1975−84 (year8594=0).
The rate for females (female=1):

$$\exp(\beta_0 + \beta_1 + \beta_{10} + \beta_{14})$$

The rate for males (female=0):

$$\exp(\beta_0 + \beta_1 + \beta_{10})$$

The mortality rate ratio of females compared to males will be:

$$\frac{\exp(\beta_0+\beta_1+\beta_{10}+\beta_{14})}{\exp(\beta_0+\beta_1+\beta_{10})} = \frac{\exp(\beta_0)\exp(\beta_1)\exp(\beta_{10})\exp(\beta_{14})}{\exp(\beta_0)\exp(\beta_1)\exp(\beta_{10})} = \exp(\beta_{14}) = 0.6031338$$

Similarly, for patients diagnosed during 1985−94 (year8594=1).
The rate for females (female=1):

$$\exp(\beta_0+\beta_1+\beta_{10}+\beta_{13}+\beta_{14}+\beta_{15})$$

The rate for males (female=0):

$$\exp(\beta_0+\beta_1+\beta_{10}+\beta_{13})$$

The mortality rate ratio of females compared to males will be:

$$\begin{aligned}\frac{\exp(\beta_0+\beta_1+\beta_{10}+\beta_{13}+\beta_{14}+\beta_{15})}{\exp(\beta_0+\beta_1+\beta_{10}+\beta_{13})} &= \frac{\exp(\beta_0)\exp(\beta_1)\exp(\beta_{10})\exp(\beta_{13})\exp(\beta_{14})\exp(\beta_{15})}{\exp(\beta_0)\exp(\beta_1)\exp(\beta_{10})\exp(\beta_{13})}\\ &= \exp(\beta_{14})\exp(\beta_{15}) = 0.6031338 \times 0.9437245 = 0.569192\end{aligned}$$

So, the hazard ratios are the same when comparing females to males within this follow-up year and age group, as it was within the reference follow-up year and reference age group. And you should be able to see from above that it will be the case irrespective of which age group and follow-up year you compare within, since they will always cancel out as long as we are comparing within the same category. This is the case because we don't have an interaction between sex and age and between sex and follow-up year.
For completeness (although this was not part of the exercise), we can also make comparisons across different levels for the other covariates. For example, what is the hazard ratio comparing females within the first year of follow-up aged 45-59 at diagnosis to males within the first year of follow-up aged 75+ at diagnosis, in each of the calendar periods. The same model is here written again:

$$\begin{aligned}\lambda = \exp(&\beta_0+\beta_1\mathrm{fu}_{1-2}+\beta_2\mathrm{fu}_{2-3}+\beta_3\mathrm{fu}_{3-4}+\beta_4\mathrm{fu}_{4-5}+\beta_5\mathrm{fu}_{5-6}+\beta_6\mathrm{fu}_{6-7}+\beta_7\mathrm{fu}_{7-8}+\beta_8\mathrm{fu}_{8-9}+\beta_9\mathrm{fu}_{9-10}\\ &+\beta_{10}\mathrm{age}_{45-59}+\beta_{11}\mathrm{age}_{60-74}+\beta_{12}\mathrm{age}_{75+}+\beta_{13}\mathrm{year8594}+\beta_{14}\mathrm{female}+\beta_{15}(\mathrm{year8594}\times\mathrm{female}))\end{aligned}$$

Let's now obtain the mortality rate ratio for females at the first year of follow-up (reference level for fu) and at age-group 45-59 ($\mathrm{age}_{45-59}=1$) compared to males at the first year of follow-up (reference level of fu) and at age-group 75+ ($\mathrm{age}_{75+}=1$).
First for patients diagnosed during 1975−84 (year8594=0).
The rate for females, diagnosed aged 45-59 in first calendar period and within the first year of follow-up:

$$\exp(\beta_0+\beta_{10}+\beta_{14})$$

The rate for males, diagnosed aged 75+ in first calendar period and within the first year of follow-up:

$$\exp(\beta_0+\beta_{12})$$

The mortality rate ratio will be:

$$\begin{aligned}\frac{\exp(\beta_0+\beta_{10}+\beta_{14})}{\exp(\beta_0+\beta_{12})} = \frac{\exp(\beta_0)\exp(\beta_{10})\exp(\beta_{14})}{\exp(\beta_0)\exp(\beta_{12})} &= \frac{\exp(\beta_{10})\exp(\beta_{14})}{\exp(\beta_{12})}\\ &= (1.326709 \times 0.6031338)/3.399539 = 0.23537987\end{aligned}$$

Similarly, for patients diagnosed during 1985−94 (year8594=1).
The rate for females (female=1):

$$\exp(\beta_0 + \beta_{10} + \beta_{13} + \beta_{14} + \beta_{15})$$

The rate for males (female=0):

$$\exp(\beta_0 + \beta_{12} + \beta_{13})$$

The mortality rate ratio of females compared to males will be:

$$\begin{aligned}
\frac{\exp(\beta_0 + \beta_{10} + \beta_{13} + \beta_{14} + \beta_{15})}{\exp(\beta_0 + \beta_{12} + \beta_{13})} &= \frac{\exp(\beta_0)\exp(\beta_{10})\exp(\beta_{13})\exp(\beta_{14})\exp(\beta_{15})}{\exp(\beta_0)\exp(\beta_{12})\exp(\beta_{13})} \\
&= \frac{\exp(\beta_{10})\exp(\beta_{14})\exp(\beta_{15})}{\exp(\beta_{12})} \\
&= (1.326709 \times 0.6031338 \times 0.9437245)/3.399539 = 0.22213375
\end{aligned}$$

If we instead use the other type of parameterisation, as suggested in 111k iv) (which is the same as in 111k iii) where you are asked to create your own variables instead of using the Stata # syntax), the output and model formulation is as shown below:

```
. streg i.fu i.agegrp i.year8594 i.year8594#i.sex, dist(exp)

------------------------------------------------------------------------------------------
                    _t | Haz. Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-----------------------+------------------------------------------------------------------
                    fu |
                    1  |   3.554795   .4831838     9.33   0.000     2.723425    4.639955
                    2  |   3.693547   .5099324     9.46   0.000     2.817906    4.841287
                    3  |   2.932013   .4288725     7.35   0.000     2.201195    3.905468
                    4  |   2.447604   .3808316     5.75   0.000     1.804262    3.320341
                    5  |    2.25602   .3692772     4.97   0.000     1.636868    3.109367
                    6  |   1.797325   .3227558     3.26   0.001     1.264071    2.555534
                    7  |   1.288401    .267454     1.22   0.222     .8577355    1.935301
                    8  |   1.439152   .3023187     1.73   0.083     .9534478    2.172282
                    9  |   .7958958    .221615    -0.82   0.412     .4611492    1.373634
                       |
                agegrp |
                45-59  |   1.326709   .1249663     3.00   0.003     1.103059    1.595705
                60-74  |   1.861131   .1691561     6.83   0.000     1.557443    2.224035
                  75+  |   3.399539   .3550374    11.72   0.000     2.770277    4.171737
                       |
              year8594 |
        Diagnosed 85-94  |   .7414351   .0655414    -3.38   0.001     .6234888    .8816936
                       |
          year8594#sex |
Diagnosed 75-84#Female  |   .6031338   .0531555    -5.74   0.000     .5074526     .716856
Diagnosed 85-94#Female  |   .5691922    .055267    -5.80   0.000     .4705541    .6885069
                       |
                 _cons |   .0125379     .00183   -30.00   0.000     .0094185    .0166904
------------------------------------------------------------------------------------------
```

$$\begin{aligned}\ln(\lambda) = \beta_0 &+ \beta_1 \text{fu}_{1-2} + \beta_2 \text{fu}_{2-3} + \beta_3 \text{fu}_{3-4} + \beta_4 \text{fu}_{4-5} + \beta_5 \text{fu}_{5-6} + \beta_6 \text{fu}_{6-7} + \beta_7 \text{fu}_{7-8} + \beta_8 \text{fu}_{8-9} + \beta_9 \text{fu}_{9-10} \\ &+ \beta_{10} \text{age}_{45-59} + \beta_{11} \text{age}_{60-74} + \beta_{12} \text{age}_{75+} + \beta_{13} \text{year8594} \\ &+ \beta_{14}(\text{year7584} \times \text{female}) + \beta_{15}(\text{year8594} \times \text{female})\end{aligned}$$

On the rate scale this can be written as:

$$\begin{aligned}\lambda = \exp(\beta_0 &+ \beta_1 \text{fu}_{1-2} + \beta_2 \text{fu}_{2-3} + \beta_3 \text{fu}_{3-4} + \beta_4 \text{fu}_{4-5} + \beta_5 \text{fu}_{5-6} + \beta_6 \text{fu}_{6-7} + \beta_7 \text{fu}_{7-8} + \beta_8 \text{fu}_{8-9} + \beta_9 \text{fu}_{9-10} \\ &+ \beta_{10} \text{age}_{45-59} + \beta_{11} \text{age}_{60-74} + \beta_{12} \text{age}_{75+} + \beta_{13} \text{year8594} \\ &+ \beta_{14}(\text{year7584} \times \text{female}) + \beta_{15}(\text{year8594} \times \text{female}))\end{aligned}$$

Note the inclusion of year7584 in one of the interaction terms in the models above. Keep in mind that even though the parametrization is different to the model fitted in 111 j), the underlying model is still the same.
We want to know what the hazard ratio is for females compared to males, in both calendar periods. The hazard ratios will be the same within all ages and follow-up years, as long as

we compare males and females within the same age and follow-up year, since we don't have any interaction between those covariates (if this is not clear, write out the rates for males and emales for different follow-up years and ages). For simplicity, let's estimate the rates within the first year of follow-up and the youngest age group.
Let's start with the first calendar period, so those diagnosed in the years 1975-84:
The rate for females (female=1):

$$\exp(\beta_0 + \beta_{14})$$

The rate for males (female=0):

$$\exp(\beta_0)$$

The mortality rate ratio of females compared to males will be:

$$\frac{\exp(\beta_0 + \beta_{14})}{\exp(\beta_0)} = \frac{\exp(\beta_0)\exp(\beta_{14})}{\exp(\beta_0)} = \exp(\beta_{14}) = 0.6031338$$

Similarly, for patients diagnosed during 1985−94 (year8594=1).
The rate for females (female=1):

$$\exp(\beta_0 + \beta_{13} + \beta_{15})$$

The rate for males (female=0):

$$\exp(\beta_0 + \beta_{13})$$

The mortality rate ratio of females compared to males will be:

$$\frac{\exp(\beta_0 + \beta_{13} + \beta_{15})}{\exp(\beta_0 + \beta_{13})} = \frac{\exp(\beta_0)\exp(\beta_{13})\exp(\beta_{15})}{\exp(\beta_0))\exp(\beta_{13})} = \exp(\beta_{15}) = 0.569192$$

These are the same hazard ratios as we got from the parameterisation used in 111k).