## Stata tips #3

***Adapted from previous TAs’ notes and tips***

This week:

Using cs, ir commands (REVIEW)

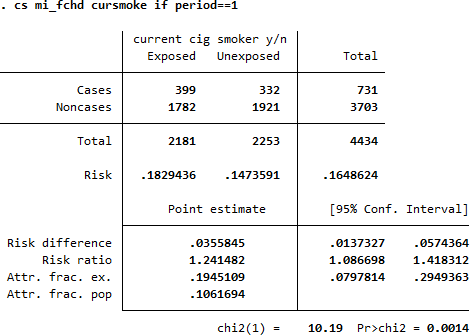
# cs, and ir:

**Cohort study command**

The command *cs* is used where you are analyzing cohort study data where there is equal follow-up time for each subject; this can also be useful with cross-sectional data. This is based on cumulative incidence; risk is calculated based on the proportion of cases. The basic syntax is

# cs variable1 variable2

where the first variable represents the case or outcome of interest, and the second variable is the exposure or predictor variable. To practice this in the Framingham data set, I’ll use the wide format of the Framingham dataset.

The *cs* command can be run with two binary variables and this command is really just an extension of the tab command. We’ll analyze a couple of variables that are already in binary form by comparing *mi\_fchd1* and *cursmoke1.*

You can see that the output at the top of the table is very similar to the tab command, except that it is more descriptive from an epidemiological perspective. Here, the number of cases is the number of subjects who had an MI and the exposed column denotes those who were smokers (based on how the variable was coded). In addition, the chi-squared value and P-value are the exact same as the tab command.

As shown in the bottom half of the table above, the cs command provides additional valuable information beyond what you get from the tab command. First, under the “Total” row is the “Risk” row that provides the proportion of MI subjects in the exposed and the unexposed group. This comes from dividing the “Cases” row by the “Total” row, and this can be seen by the calculation below:

. di 399/2181

.1829436

. di 332/2253 .14735908

The following four rows all provide additional information that compares the risk of MI for those who are exposed (smokers) to those who were unexposed (nonsmokers).

What is the **risk difference**?

. di 0.183-0.147

.036

What is the **risk ratio**?

. di 0.183/0.147 1.244898

What about the **Attr. frac. ex. and pop.**?

. di (0.183-0.147)/0.183

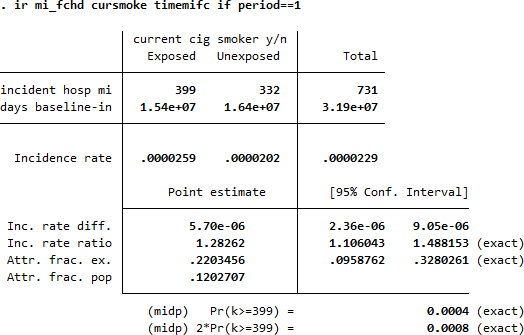
.19672131

. di (0.165-0.147)/0.164

.1097561

# Incidence ratio command

This command is conceptually similar to *cs* but takes into accountperson-time information. This means that you must enter a time-related variable to enable the computation. The syntax involves listing the outcome, exposure and time variables, in that order:



Now the second row in the table is the total amount of person-time in the exposed group and in the unexposed group. Dividing the number of MI cases by the total amount of time observed for each group ten gives you the incidence rate. The next two rows are similar to the cs command: the **incidence rate difference and the incidence rate ratio**.