WWS 508b
Precept 5

John Palmer

March 23, 2010

transforming variables in Stata

difference between values and labels

Let's say we want to work with the `rincom98` variable in the GSS data. We codebook it and get this:

```
-------------------------------------------------------------------------------------------------
rincom98                                                                       respondents income
-------------------------------------------------------------------------------------------------

          type:  numeric (byte)
         label:  rincom98

         range:  [1,98]                       units:  1
  unique values:  25                     missing .:  963/2812

      examples:  13    $20000 to 22499
                 18    $40000 to 49999
                 23    $110 000 over
                 .
```

What should we type to create a new dummy variable indicating incomes of $20,000 and over?

transforming variables in Stata

difference between values and labels

Let's say we want to work with the `rincom98` variable in the GSS data. We codebook it and get this:

```
------------------------------------------------------------------------------------------------------
rincom98                                                                            respondents income
------------------------------------------------------------------------------------------------------

              type:  numeric (byte)
             label:  rincom98

             range:  [1,98]                        units:  1
     unique values:  25                         missing .:  963/2812

          examples:  13     $20000 to 22499
                     18     $40000 to 49999
                     23     $110 000 over
                     .
```

What should we type to create a new dummy variable indicating incomes of \$20,000 and over?

```
gen dummy = 1 if rincom98 >= 13 & rincom98 <= 98
replace dummy = 0 if rincom98 < 13
```

Let's say you want to fit the model

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x$$

but you want your intercept term $(\hat{\beta}_0)$ to be equal to zero. In other words, you want your regression line to run through the origin. How can you do this?

Let's say you want to fit the model

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x$$

but you want your intercept term ($\hat{\beta}_0$) to be equal to zero. In other words, you want your regression line to run through the origin. How can you do this?

In Stata, just add the option `nocontant` to the regression command. E.g.:

```
reg y x, noconstant
```

Bayes' Rule vs. standard conditional probability

Can Question 1 be done using the conditional probability rules from the "Introduction/Review" handout?

Yes. You can use Bayes' Rule if you know it and want to, but you should be able to figure out the answer based on the handout. If you're curious, Bayes' Rule is:

$$\Pr\{A|B\} = \frac{\Pr\{B|A\}\Pr\{A\}}{\Pr\{B\}}$$

Questions about testing proportions

Review solution to Question C9 in Problem set 1

Questions about graphing

Review do file for Problem Set 3