

brain2.sas: Explanation of SAS code

Goal:

demonstrate ways to include quadratic and cross-product terms in a multiple regression

This uses the brain data set, in brain.txt. The first few lines read the data and compute the log transform of each variable.

Remember: if you use proc import to read a .xlsx or .csv file, you need a separate data step, with a set command, to compute new variables.

Storing a quadratic or cross product term in a new variable:

```
logbody2 = logbody**2
```

Much of this is familiar. The ** operator is the “raise to a power” operator. **2 computes x squared. x**3 would compute x cubed. You could also write logbody*logbody.

Crossproduct terms are computed and saved in the obvious way.

The quadratic or crossproduct variable is used in a multiple regression just like any other variable. This is illustrated by the first proc glm; .

If you want to use proc reg to fit the regression with quadratic or cross product term, you must create new variables. **Computing a quadratic term “on the fly”:** logbody*logbody
If you use proc glm, you do not need to create new variables first. You can specify them in the model statement. You can’t use the power operator, but you can compute the product in the model statement.

Again, you can compute terms “on the fly” in proc glm but not in proc reg.

Computing cross product terms “on the fly”: logbody*loglitter

Done in the same way as quadratic terms. Again, works in proc glm;, not in proc reg.

Cross products of continuous and class variables:

Nothing new - Write the cross product in the model formula. There, SAS doesn’t care whether the components are two continuous variables, a continuous and a class, or two class variables.