

LAB 5: Backpain Data Analysis

Recall the backpain dataset from Labs 3 and 4 with the following variables:

Doctor – Doctor ID number.

Cost – Cost of treatment in dollars.

Logcost – Logarithm of cost.

Actlim – Number of days in the past six months in which activity has been limited.

Undrstnd – Did the patient understand the doctor's advice (1=yes, 0=no)

Age – Age of the patient.

Educ – Education of the patient (0 means <12 years, 1 means 13-16 years, 2 means > 16 years education).

Thoraic – Whether the back pain was cervical/thoraic or other (1=yes, 0=no).

Pracstyl – Practice style of the doctor (0=low, 1=medium, 2=high frequency of prescription of medicines and hospitalization for treatment of back pain).

Remember that the outcomes are Cost or Logcost, Actlim and Undrstnd and that the primary predictor of interest is the practice style of the physician. You need to do the following things for each outcome:

1. Decide on an analysis method.
2. Decide on which variables you need to include in the model. Is there evidence of interactions? Be careful when handling age to check the form of the relationship.
3. Provide an interpretation of the effect of practice style.

Warning: these are just to get you started and are not final models.

```
proc mixed data=work.backpain;
  class doctor educ pracstyl ;
  model cost=educ pracstyl age thoraic/ddfm=kenrog;
  random doctor;
run;
proc genmod data=work.backpain;
  class doctor educ pracstyl ;
  model undrstnd=educ pracstyl age thoraic/dist=bin;
  repeated subject=doctor/corrw type=exch;
run;
proc nlmixed data=work.backpain;
  parms b0=0 b1=0 sig1=0.1;
  eta=b0+b1*age+u;
  p=1/(1+exp(-eta));
  model undrstnd ~ binary(p);
  random u~Normal(0,sig1) subject=doctor;
run;
```

(See Lab 4 for sample R code).