Solution of week 2

Michele De Vita

March 12, 2018

1. *drop if !inlist(schoolid, 1224, 1288, 1296, 1308, 1317)*
   keep mathach minority schoolid ses

2. - cmd for continuos variables: *tabstat ses mathach, stat(mean sd)*
   - cmd for categorical: *table minority*
   - continuos vears by schoolid: *tabstat ses mathach, stat(mean sd) by(schoolid)*

3. *histogram mathach*
   graph box mathach

4. *scatter mathach ses*
   twoway scatter mathach ses, by(schoolid)

5. *regress mathach ses*
   Parameters interpretation:

   b) With intercept equal to 11.45652 we can say that if we have $ses = 0$ we have a math achievement equal to 11.45652
   The coefficient β_1 before *ses* say that if we increment ses (socio economic status) by one we have an increment of mathach by 3.306963
   The residual ϵ_i $N(0, \sigma^2)$ with $\sigma = 6.4708$ and $\sigma^2 = 41.8712$ tell us about the average deviation from the regression line.

   c) Since the p-value is less that 0.05 we reject the null hypothesis $H_0 = \beta_2 = 0$

6. *predict yhat*

7. *twoway (scatter mathach ses) (lfit yhat ses)*
   twoway (scatter mathach ses) (lfit yhat ses), by(schoolid)
   twoway (scatter mathach ses) (lfit yhat ses) (lfit mathach ses), by(schoolid)

   Comment about differences between mean mathach, ses and yhat, ses: For the school 1224 and 1288 the lines respectively $\hat{y}$ and are very similar, while the other three are a little bit different

8. a) *tabulate schoolid, generate(schoolid_)*

b) *regress mathach schoolid_2 schoolid_3 schoolid_4 schoolid_5* The coefficient before the schoolids dummy variables say the differences between the school with $schoolid = 1224$ and the others. The coefficients say that:

- $\beta_1 = 3.795353$ say that the school with $schoolid = 1228$ has an average mathach superior by 3.795353 from the school with $schoolid = 1224$.
- $\beta_2 = -2.079489$ say that the school with $schoolid = 1296$ has a mean mathach lower by 2.079489 respect to school with $schoolid = 1224$
- .. and so on with the other two coefficient and respectively the other two id of the schools

c) *testparm schoolid_∗* We reject the null hypothesis of all dummy variables are 0 because the p-value is 0.0015, so less than 0.025

9. *regress mathach c.ses##i.schoolid_∗*