Okay, I am working on the example you posted in week 1 and this is riveting.   
  
I am to the point in code where:

**proc** **reg** data=example1;

model mother\_ed=education experience ability;

output out=resid1 r=r1;

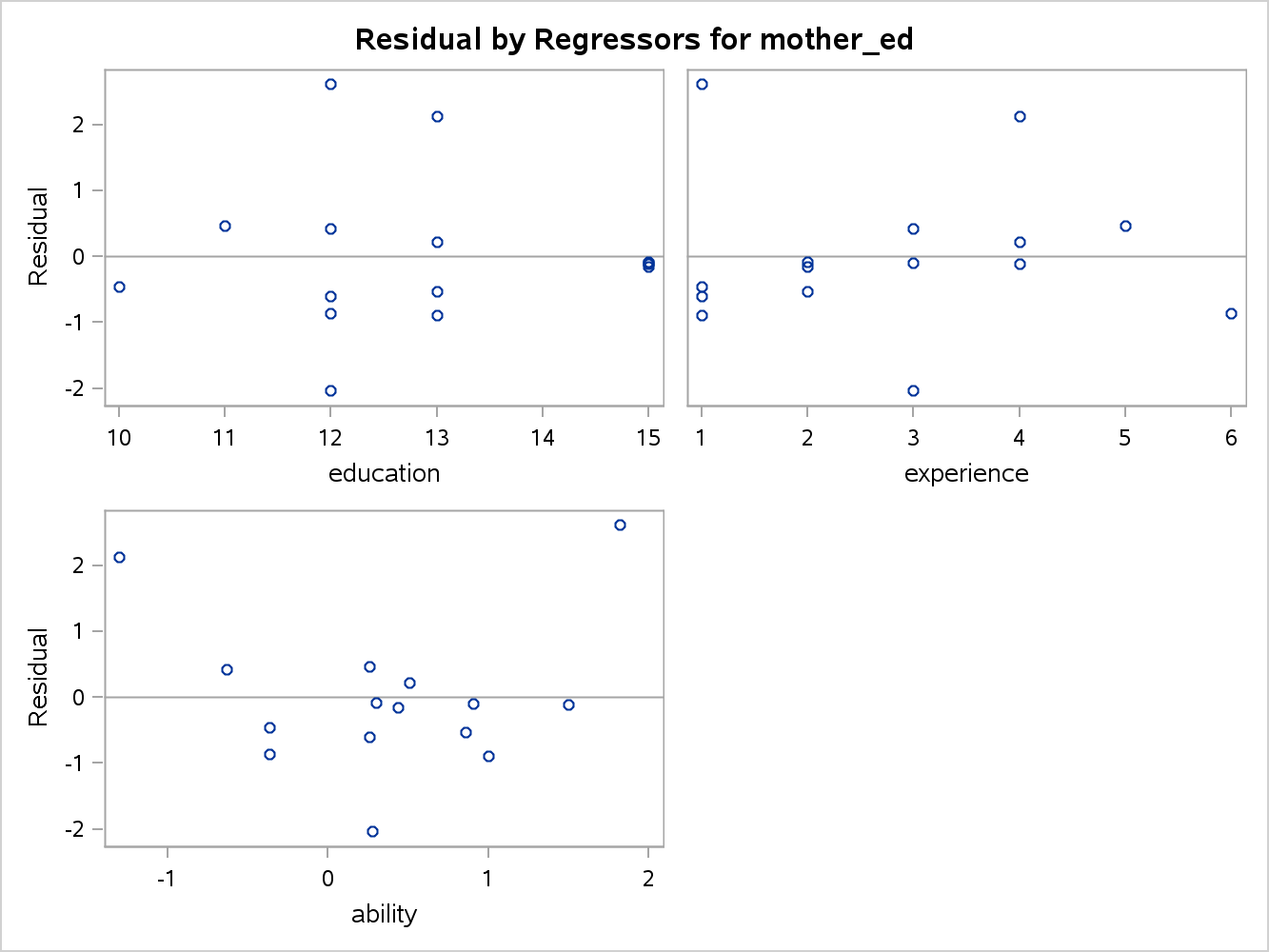
**run**;

**proc** **means** data=resid1;

var r1;

How is the mean for resid1 calculated?

Is it the mean of all these residual points?



Once this question is answered it will be answered for the other residuals as well.

Also, these variables - education experience ability; found the foundation for the model of which we are building our residuals.

Why are we having these in the model? They are not predictive?

Perhaps, this is just for the sake of the example. In practicality. we could reverse the situation and have education mom/dad and sibling be the core and use the residuals to form the base for education, experience, and ability.

I just have to know, why did we not include an intercept?

It changed everything.

I understand that the intercept is the average for each column, but why does it affect everything so much? The r-score especially was affected.

| **Parameter Estimates** | | | | | | |
| --- | --- | --- | --- | --- | --- | --- |
| **Variable** | **Label** | **DF** | **Parameter Estimate** | **Standard Error** | **t Value** | **Pr > |t|** |
| **Intercept** | Intercept | 1 | 1.66364 | 0.55831 | 2.98 | 0.0176 |
| **education** |  | 1 | 0.01454 | 0.04425 | 0.33 | 0.7509 |
| **experience** |  | 1 | 0.07103 | 0.04336 | 1.64 | 0.1400 |
| **ability** |  | 1 | 0.02662 | 0.08946 | 0.30 | 0.7737 |
| **r1** | Residual | 1 | 0.10163 | 0.07018 | 1.45 | 0.1856 |
| **r2** | Residual | 1 | 0.00164 | 0.04465 | 0.04 | 0.9715 |
| **r3** | Residual | 1 | 0.05917 | 0.06902 | 0.86 | 0.4162 |

| **Parameter Estimates** | | | | | |
| --- | --- | --- | --- | --- | --- |
| **Variable** | **DF** | **Parameter Estimate** | **Standard Error** | **t Value** | **Pr > |t|** |
| **Intercept** | 1 | 0.04900 | 0.94881 | 0.05 | 0.9601 |
| **education** | 1 | 0.02582 | 0.04469 | 0.58 | 0.5793 |
| **experience** | 1 | 0.10339 | 0.04735 | 2.18 | 0.0605 |
| **ability** | 1 | 0.03074 | 0.12120 | 0.25 | 0.8062 |
| **mother\_ed** | 1 | 0.10163 | 0.07018 | 1.45 | 0.1856 |
| **father\_ed** | 1 | 0.00164 | 0.04465 | 0.04 | 0.9715 |
| **siblings** | 1 | 0.05917 | 0.06902 | 0.86 | 0.4162 |

While the residual coefficients are the same, the intercept is bigger with the residual model, and the pvalues and tvalues are different as well. What are your thoughts on this?