The data file is a subsample from the 1982 High School and Beyond Survey and is used extensively in *Hierarchical Linear Models* by Raudenbush and Bryk. The data file consists of 7185 students nested in 160 schools. The outcome variable of interest is student-level math achievement score (**MATHACH**). Variable **SES** is social-economic-status of a student and therefore is a student-level variable. Variable **MEANSES** is the group mean of **SES** and therefore is a school-level variable. Both **SES** and **MEANSES** are centered at the grand mean (they both have means of 0). Variable **SECTOR** is an indicator variable indicating if a school is public or catholic and is therefore a school-level variable. There are 90 public schools (SECTOR=0) and 70 catholic schools (SECTOR=1) in the sample.

**proc** **print** data=MATHACH (obs=**5**);

**run**;

| **Obs** | **SCHOOL** | **MINORITY** | **FEMALE** | **SES** | **MATHACH** | **SIZE** | **SECTOR** | **PRACAD** |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **1** | 1224 | 0 | 1 | -1.528 | 5.876 | 842 | 0 | 0.35 |
| **2** | 1224 | 0 | 1 | -0.588 | 19.708 | 842 | 0 | 0.35 |
| **3** | 1224 | 0 | 0 | -0.528 | 20.349 | 842 | 0 | 0.35 |
| **4** | 1224 | 0 | 0 | -0.668 | 8.781 | 842 | 0 | 0.35 |
| **5** | 1224 | 0 | 0 | -0.158 | 17.898 | 842 | 0 | 0.35 |

| **Obs** | **DISCLIM** | **HIMINTY** | **MEANSES** |
| --- | --- | --- | --- |
| **1** | 1.597 | 0 | -0.428 |
| **2** | 1.597 | 0 | -0.428 |
| **3** | 1.597 | 0 | -0.428 |
| **4** | 1.597 | 0 | -0.428 |
| **5** | 1.597 | 0 | -0.428 |

**Model 1:**Unconditional Means Model

This model is referred as a one-way ANOVA with random effects and is the simplest possible random effect linear model and is discussed in detail by Raudenbush and Bryk. The motivation for this model is the question on how much schools vary in their mean mathematics achievement. In terms of regression equations, we have the following, where rij ~ N(0, σ2) and u0j~ N(0, τ2),

**proc** **mixed** data = MATHACH covtest noclprint;

class school;

model mathach = / solution;

random intercept / subject = school;

**run**;

| **Model Information** | |
| --- | --- |
| **Data Set** | WORK.MATHACH |
| **Dependent Variable** | MATHACH |
| **Covariance Structure** | Variance Components |
| **Subject Effect** | SCHOOL |
| **Estimation Method** | REML |
| **Residual Variance Method** | Profile |
| **Fixed Effects SE Method** | Model-Based |
| **Degrees of Freedom Method** | Containment |

| **Dimensions** | |
| --- | --- |
| **Covariance Parameters** | 2 |
| **Columns in X** | 1 |
| **Columns in Z per Subject** | 1 |
| **Subjects** | 160 |
| **Max Obs per Subject** | 67 |

| **Number of Observations** | |
| --- | --- |
| **Number of Observations Read** | 7185 |
| **Number of Observations Used** | 7185 |
| **Number of Observations Not Used** | 0 |

| **Iteration History** | | | |
| --- | --- | --- | --- |
| **Iteration** | **Evaluations** | **-2 Res Log Like** | **Criterion** |
| **0** | 1 | 48102.91726234 |  |
| **1** | 2 | 47116.81230623 | 0.00000109 |
| **2** | 1 | 47116.79350024 | 0.00000000 |

|  |
| --- |
| Convergence criteria met. |

| **Covariance Parameter Estimates** | | | | | |
| --- | --- | --- | --- | --- | --- |
| **Cov Parm** | **Subject** | **Estimate** | **Standard Error** | **Z Value** | **Pr > Z** |
| **Intercept** | SCHOOL | 8.6097 | 1.0778 | 7.99 | <.0001 |
| **Residual** |  | 39.1487 | 0.6607 | 59.26 | <.0001 |

| **Fit Statistics** | |
| --- | --- |
| **-2 Res Log Likelihood** | 47116.8 |
| **AIC (Smaller is Better)** | 47120.8 |
| **AICC (Smaller is Better)** | 47120.8 |
| **BIC (Smaller is Better)** | 47126.9 |

| **Solution for Fixed Effects** | | | | | |
| --- | --- | --- | --- | --- | --- |
| **Effect** | **Estimate** | **Standard Error** | **DF** | **t Value** | **Pr > |t|** |
| **Intercept** | 12.6370 | 0.2443 | 159 | 51.72 | <.0001 |

**Model 2**: Including Effects of School Level (level 2) Predictors —predicting**mathach**from**meanses**

This model is referred as regression with Means-as-Outcomes by Raudenbush and Bryk. The motivation of this model is the question on if the schools with high **MEANSES** also have high math achievement. In other words, we want to understand why there is a school difference on mathematics achievement. In terms of regression equations, we have the following.

MATHACHij=  β0j+ rij

β0j=  γ00+ γ01(MEANSES) + u0j

Combining the two equations into one by substituting the level-2 equation to level-1 equation, we have

MATHACHij=  γ00+ γ01(MEANSES) + u0j+ rij

**proc** **mixed** data = MATHACH covtest noclprint;

class school;

model mathach = meanses / solution ddfm = bw;

random intercept / subject = school;

**run**;

| **Model Information** | |
| --- | --- |
| **Data Set** | WORK.MATHACH |
| **Dependent Variable** | MATHACH |
| **Covariance Structure** | Variance Components |
| **Subject Effect** | SCHOOL |
| **Estimation Method** | REML |
| **Residual Variance Method** | Profile |
| **Fixed Effects SE Method** | Model-Based |
| **Degrees of Freedom Method** | Between-Within |

| **Dimensions** | |
| --- | --- |
| **Covariance Parameters** | 2 |
| **Columns in X** | 2 |
| **Columns in Z per Subject** | 1 |
| **Subjects** | 160 |
| **Max Obs per Subject** | 67 |

| **Number of Observations** | |
| --- | --- |
| **Number of Observations Read** | 7185 |
| **Number of Observations Used** | 7185 |
| **Number of Observations Not Used** | 0 |

| **Iteration History** | | | |
| --- | --- | --- | --- |
| **Iteration** | **Evaluations** | **-2 Res Log Like** | **Criterion** |
| **0** | 1 | 47201.23573408 |  |
| **1** | 2 | 46961.28490236 | 0.00000000 |

|  |
| --- |
| Convergence criteria met. |

| **Covariance Parameter Estimates** | | | | | |
| --- | --- | --- | --- | --- | --- |
| **Cov Parm** | **Subject** | **Estimate** | **Standard Error** | **Z Value** | **Pr > Z** |
| **Intercept** | SCHOOL | 2.6357 | 0.4036 | 6.53 | <.0001 |
| **Residual** |  | 39.1578 | 0.6608 | 59.26 | <.0001 |

| **Fit Statistics** | |
| --- | --- |
| **-2 Res Log Likelihood** | 46961.3 |
| **AIC (Smaller is Better)** | 46965.3 |
| **AICC (Smaller is Better)** | 46965.3 |
| **BIC (Smaller is Better)** | 46971.4 |

| **Solution for Fixed Effects** | | | | | |
| --- | --- | --- | --- | --- | --- |
| **Effect** | **Estimate** | **Standard Error** | **DF** | **t Value** | **Pr > |t|** |
| **Intercept** | 12.6495 | 0.1492 | 158 | 84.77 | <.0001 |
| **MEANSES** | 5.8635 | 0.3613 | 158 | 16.23 | <.0001 |

| **Type 3 Tests of Fixed Effects** | | | | |
| --- | --- | --- | --- | --- |
| **Effect** | **Num DF** | **Den DF** | **F Value** | **Pr > F** |
| **MEANSES** | 1 | 158 | 263.37 | <.0001 |

**Model 3:**Including Effects of Student-Level Predictors–predicting**mathach**fromcentered student-level ses, **cses**

This model is referred as a random-coefficient model by Raudenbush and Bryk. Pretend that we run regression of **mathach** on centered **ses** on each school, that is we are going to run 160 regressions.

1. What would be the average of the 160 regression equations (both intercept and slope)?
2. How much do the regression equations vary from school to school?
3. What is the correlation between the intercepts and slopes?

These are some of the questions that motivates the following model.

MATHACHij=  β0j+ β1j (SES – MEANSES) + rij

β0j=  γ00 + u0j

β1j=  γ10 + u1j

Combining the two equations into one by substituting the level-2 equation to level-1 equation, we have

MATHACHij= γ00+ γ10(SES – MEANSES) + u0j+  u1j(SES – MEANSES) + rij

**data** MATHACH2;

set MATHACH;

cses = ses - meanses;

**run**;

**proc** **mixed** data = MATHACH2 noclprint covtest noitprint;

class school;

model mathach = cses / solution ddfm = bw notest;

random intercept cses / subject = school type = un gcorr;

**run**;

| **Model Information** | |
| --- | --- |
| **Data Set** | WORK.MATHACH2 |
| **Dependent Variable** | MATHACH |
| **Covariance Structure** | Unstructured |
| **Subject Effect** | SCHOOL |
| **Estimation Method** | REML |
| **Residual Variance Method** | Profile |
| **Fixed Effects SE Method** | Model-Based |
| **Degrees of Freedom Method** | Between-Within |

| **Dimensions** | |
| --- | --- |
| **Covariance Parameters** | 4 |
| **Columns in X** | 2 |
| **Columns in Z per Subject** | 2 |
| **Subjects** | 160 |
| **Max Obs per Subject** | 67 |

| **Number of Observations** | |
| --- | --- |
| **Number of Observations Read** | 7185 |
| **Number of Observations Used** | 7185 |
| **Number of Observations Not Used** | 0 |

| **Estimated G Correlation Matrix** | | | | |
| --- | --- | --- | --- | --- |
| **Row** | **Effect** | **SCHOOL** | **Col1** | **Col2** |
| **1** | Intercept | 1224 | 1.0000 | 0.02068 |
| **2** | cses | 1224 | 0.02068 | 1.0000 |

| **Covariance Parameter Estimates** | | | | | |
| --- | --- | --- | --- | --- | --- |
| **Cov Parm** | **Subject** | **Estimate** | **Standard Error** | **Z Value** | **Pr Z** |
| **UN(1,1)** | SCHOOL | 8.6769 | 1.0786 | 8.04 | <.0001 |
| **UN(2,1)** | SCHOOL | 0.05075 | 0.4062 | 0.12 | 0.9006 |
| **UN(2,2)** | SCHOOL | 0.6940 | 0.2808 | 2.47 | 0.0067 |
| **Residual** |  | 36.7006 | 0.6258 | 58.65 | <.0001 |

| **Fit Statistics** | |
| --- | --- |
| **-2 Res Log Likelihood** | 46714.2 |
| **AIC (Smaller is Better)** | 46722.2 |
| **AICC (Smaller is Better)** | 46722.2 |
| **BIC (Smaller is Better)** | 46734.5 |

| **Null Model Likelihood Ratio Test** | | |
| --- | --- | --- |
| **DF** | **Chi-Square** | **Pr > ChiSq** |
| 3 | 1065.70 | <.0001 |

| **Solution for Fixed Effects** | | | | | |
| --- | --- | --- | --- | --- | --- |
| **Effect** | **Estimate** | **Standard Error** | **DF** | **t Value** | **Pr > |t|** |
| **Intercept** | 12.6493 | 0.2445 | 159 | 51.75 | <.0001 |
| **cses** | 2.1932 | 0.1283 | 7024 | 17.10 | <.0001 |

**Model 4:**Including Both Level-1 and Level-2 Predictors –predicting**mathach**from**meanses, sector, cses** and the cross-level interaction of**meanses**and **sector** with **cses**

This model is referred as an intercepts and slopes-as-outcomes model by Raudenbush and Bryk. We have examined the variability of the regression equations across schools. Now we will build an explanatory model to account for the variability. That is we want to model the following:

MATHACHij=  β0j+ β1j (SES – MEANSES) + rij

β0j=  γ00 + γ01(MEANSES) + γ02(SECTOR) + u0j

β1j=  γ10 + γ11(MEANSES) + γ12(SECTOR) + u1j

Combining the two equations into one by substituting the level-2 equation to level-1 equation, we have

MATHACHij=  γ00 + γ01(MEANSES) + γ02(SECTOR) + γ10 (SES – MEANSES) + γ11(MEANSES)\* (SES – MEANSES) +  γ12(SECTOR)\* (SES – MEANSES) + u0j  +u1j(SES-MEANSES) +  rij

The questions that we are interested in are:

1. Do MEANSES and SECTOR significantly predict the intercept?
2. Do MEANSES and SECTOR significantly predict the within-school slopes?
3. How much variation in the intercepts and the slopes is explained by MEANSES and SECTOR?

**proc** **mixed** data = MATHACH2 noclprint covtest noitprint;

class school;

model mathach = meanses sector cses meanses\*cses sector\*cses

/ solution ddfm = bw notest;

random intercept cses / subject = school type = un;

**run**;

| **Model Information** | |
| --- | --- |
| **Data Set** | WORK.MATHACH2 |
| **Dependent Variable** | MATHACH |
| **Covariance Structure** | Unstructured |
| **Subject Effect** | SCHOOL |
| **Estimation Method** | REML |
| **Residual Variance Method** | Profile |
| **Fixed Effects SE Method** | Model-Based |
| **Degrees of Freedom Method** | Between-Within |

| **Dimensions** | |
| --- | --- |
| **Covariance Parameters** | 4 |
| **Columns in X** | 6 |
| **Columns in Z per Subject** | 2 |
| **Subjects** | 160 |
| **Max Obs per Subject** | 67 |

| **Number of Observations** | |
| --- | --- |
| **Number of Observations Read** | 7185 |
| **Number of Observations Used** | 7185 |
| **Number of Observations Not Used** | 0 |

| **Covariance Parameter Estimates** | | | | | |
| --- | --- | --- | --- | --- | --- |
| **Cov Parm** | **Subject** | **Estimate** | **Standard Error** | **Z Value** | **Pr Z** |
| **UN(1,1)** | SCHOOL | 2.3817 | 0.3717 | 6.41 | <.0001 |
| **UN(2,1)** | SCHOOL | 0.1926 | 0.2045 | 0.94 | 0.3464 |
| **UN(2,2)** | SCHOOL | 0.1014 | 0.2138 | 0.47 | 0.3177 |
| **Residual** |  | 36.7212 | 0.6261 | 58.65 | <.0001 |

| **Fit Statistics** | |
| --- | --- |
| **-2 Res Log Likelihood** | 46503.7 |
| **AIC (Smaller is Better)** | 46511.7 |
| **AICC (Smaller is Better)** | 46511.7 |
| **BIC (Smaller is Better)** | 46524.0 |

| **Null Model Likelihood Ratio Test** | | |
| --- | --- | --- |
| **DF** | **Chi-Square** | **Pr > ChiSq** |
| 3 | 220.57 | <.0001 |

| **Solution for Fixed Effects** | | | | | |
| --- | --- | --- | --- | --- | --- |
| **Effect** | **Estimate** | **Standard Error** | **DF** | **t Value** | **Pr > |t|** |
| **Intercept** | 12.1136 | 0.1988 | 157 | 60.93 | <.0001 |
| **MEANSES** | 5.3391 | 0.3693 | 157 | 14.46 | <.0001 |
| **SECTOR** | 1.2167 | 0.3064 | 157 | 3.97 | 0.0001 |
| **cses** | 2.9388 | 0.1551 | 7022 | 18.95 | <.0001 |
| **MEANSES\*cses** | 1.0389 | 0.2989 | 7022 | 3.48 | 0.0005 |
| **SECTOR\*cses** | -1.6426 | 0.2398 | 7022 | -6.85 | <.0001 |

**proc univariate data = hsbc;**

**var meanses;**

**run;**

/\*

90% 0.523

75% Q3 0.333

50% Median 0.038

25% Q1 -0.317

10% -0.579

5% -0.696

1% -1.043

0% Min -1.188

\*/

**data** toplot;

set MATHACH2;

if meanses <= -**0.317** then do;

ms = -**0.317**;

strata = "Low"; end;

else if meanses >= **0.333** then do;

ms = **0.333**;

strata = "Hig"; end;

else do; ms = **0.038**; strata = "Med" ; end;

predicted = **12.1136** + **5.3391**\*ms + **1.2167**\*sector + **2.9388**\*cses +

**1.0389**\*ms\*cses - **1.6426**\*sector\*cses;

**run**;

**proc** **sort** data = toplot;

by strata;

**run**;

goptions reset = all;

symbol1 v = none i = join c = red ;

symbol2 v = none i = join c = blue ;

axis1 order = (-**4** to **3** by **1**) minor = none label=("Group Centered SES");

axis2 order = (**0** to **22** by **2**) minor = none label=(a = **90** "Math Achievement Score");

**proc** **gplot** data = toplot;

by strata;

plot predicted\*cses = sector / vaxis = axis2 haxis = axis1;

**run**;

**quit**;



Possibly there would be two-way interaction between meanses and sector and a three way interaction between meanses, cses and sector. We can test it by adding the interaction into the model. For example,

**proc** **mixed** data = MATHACH2 noclprint covtest noitprint;

class school;

model mathach = meanses sector cses meanses\*sector

meanses\*cses sector\*cses meanses\*sector\*cses

/ solution ddfm = bw notest;

random intercept cses / subject = school type = un;

**run**;

| **Solution for Fixed Effects** | | | | | |
| --- | --- | --- | --- | --- | --- |
| **Effect** | **Estimate** | **Standard Error** | **DF** | **t Value** | **Pr > |t|** |
| **Intercept** | 12.1842 | 0.2030 | 156 | 60.01 | <.0001 |
| **MEANSES** | 5.8732 | 0.5065 | 156 | 11.60 | <.0001 |
| **SECTOR** | 1.2430 | 0.3052 | 156 | 4.07 | <.0001 |
| **cses** | 2.9513 | 0.1616 | 7021 | 18.26 | <.0001 |
| **MEANSES\*SECTOR** | -1.1276 | 0.7355 | 156 | -1.53 | 0.1273 |
| **MEANSES\*cses** | 1.1289 | 0.4232 | 7021 | 2.67 | 0.0077 |
| **SECTOR\*cses** | -1.6407 | 0.2406 | 7021 | -6.82 | <.0001 |
| **MEANSES\*SECTOR\*cses** | -0.1888 | 0.5997 | 7021 | -0.31 | 0.7528 |

Since the variance component for slopes is very small and its corresponding p-value is 0.3177. We cannot reject the hypothesis that the slopes do not differ across schools. Similarly, we cannot reject the hypothesis that the covariance between intercepts and slopes is zero. Therefore, a simpler model can be used:

**proc** **mixed** data = MATHACH2 noclprint covtest noitprint;

class school;

model mathach = meanses sector cses meanses\*cses sector\*cses / solution ddfm = bw notest;

random intercept / subject = school;

**run**;

| **Model Information** | |
| --- | --- |
| **Data Set** | WORK.MATHACH2 |
| **Dependent Variable** | MATHACH |
| **Covariance Structure** | Variance Components |
| **Subject Effect** | SCHOOL |
| **Estimation Method** | REML |
| **Residual Variance Method** | Profile |
| **Fixed Effects SE Method** | Model-Based |
| **Degrees of Freedom Method** | Between-Within |

| **Dimensions** | |
| --- | --- |
| **Covariance Parameters** | 2 |
| **Columns in X** | 6 |
| **Columns in Z per Subject** | 1 |
| **Subjects** | 160 |
| **Max Obs per Subject** | 67 |

| **Covariance Parameter Estimates** | | | | | |
| --- | --- | --- | --- | --- | --- |
| **Cov Parm** | **Subject** | **Estimate** | **Standard Error** | **Z Value** | **Pr > Z** |
| **Intercept** | SCHOOL | 2.3752 | 0.3709 | 6.40 | <.0001 |
| **Residual** |  | 36.7661 | 0.6207 | 59.24 | <.0001 |

| **Fit Statistics** | |
| --- | --- |
| **-2 Res Log Likelihood** | 46504.8 |
| **AIC (Smaller is Better)** | 46508.8 |
| **AICC (Smaller is Better)** | 46508.8 |
| **BIC (Smaller is Better)** | 46514.9 |

| **Solution for Fixed Effects** | | | | | |
| --- | --- | --- | --- | --- | --- |
| **Effect** | **Estimate** | **Standard Error** | **DF** | **t Value** | **Pr > |t|** |
| **Intercept** | 12.1138 | 0.1986 | 157 | 60.98 | <.0001 |
| **MEANSES** | 5.3429 | 0.3690 | 157 | 14.48 | <.0001 |
| **SECTOR** | 1.2146 | 0.3061 | 157 | 3.97 | 0.0001 |
| **cses** | 2.9358 | 0.1507 | 7022 | 19.48 | <.0001 |
| **MEANSES\*cses** | 1.0441 | 0.2910 | 7022 | 3.59 | 0.0003 |
| **SECTOR\*cses** | -1.6421 | 0.2331 | 7022 | -7.04 | <.0001 |

To compare the original model with this simplified one, we can compare their -2LL’s, since the fixed portion of these two models are the same.

|  |  |  |
| --- | --- | --- |
| Model | Number of parameters | -2 LL |
| restricted | 2 | 46504.8 |
| Unrestricted | 4 | 46503.7 |

Approximately, the difference in -2LL’s is a χ2 distribution with two degrees of freedom (corresponding to the difference in the number of parameters). The p-value is .577. This justifies the use of the simpler model. The SAS program is shown below.

**data** pvalue;

df = **2**; chisq = **46504.8** - **46503.7**;

pvalue = **1** - probchi(chisq, df);

**run**;

**proc** **print** data = pvalue noobs;

**run**;

| **df** | **chisq** | **pvalue** |
| --- | --- | --- |
| 2 | 1.1 | 0.57695 |