Biostat 656: Two-level Normal (Random Intercept) lab **Updated for Q&A session**

Purpose: introduce the basic two-level models and learn STATA.
Here, we will illustrate the use of two-level models for normally distributed responses. The data set used is **popular.dta**. The dataset can be downloaded using STATA command

```
use http://www.ats.ucla.edu/stat/stata/examples/mlm_ma_hox/popular.dta, clear
```

There are 5 variables, which will be used, in this dataset.
Pupil: pupil identification number
School: school identification number
Popular: the outcome variable 'popularity' (Y), measured by a self-rating scale that range from 0 (very unpopular) to 10(very popular).
Sex: the pupil sex, 0 – boy 1—girl
Texp: teacher experience in years

The data are from 2000 pupils from 100 schools, the average school size is 20 pupils. Therefore, we have pupils nested within schools, and we need to account for the possible correlation between pupils in the same school in our model.

Q1. What is the average self-rating score?

Two-stage model: subscript *j* is for the schools and *i* is for individual pupils.

The **intercept-only** model:

$$popular_{ij} = \beta_{0j} + \varepsilon_{ij}$$
$$\beta_{0j} = \gamma_{00} + \mu_{0j}$$

$$\varepsilon_{ij} \sim N(0, \sigma^2)$$
$$\mu_{0j} \sim N(0, \tau^2)$$

β_{0j}: average score for pupil in school *j*

γ_{00}: average score for a typical school (fixed effect parameter)

μ_{0j}: school-level random intercept (random effect)

This model can be fitting in xtmixed using

. xtmixed popular || school:, mle

```
Mixed-effects ML regression                     Number of obs      =      2000
Group variable: school                          Number of groups   =       100
                                                Obs per group: min =        16
                                                               avg =      20.0
                                                               max =        26
                                                Wald chi2(0)       =         .
Log likelihood = -2556.3612                     Prob > chi2        =         .

------------------------------------------------------------------------------
     popular |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _cons |   5.307603   .0950231    55.86   0.000     5.121361    5.493845
------------------------------------------------------------------------------
```

```
------------------------------------------------------------------------------
  Random-effects Parameters  |   Estimate   Std. Err.     [95% Conf. Interval]
-----------------------------+------------------------------------------------
school: Identity             |
                   sd(_cons) |   .9331053   .0684433      .8081556    1.077374
-----------------------------+------------------------------------------------
                sd(Residual) |   .7991726   .0129645      .7741624    .8249907
------------------------------------------------------------------------------
LR test vs. linear regression: chibar2(01) =  1376.81 Prob >= chibar2 = 0.0000
```

Or using xtreg (Similar results would be obtained)

```
. xtreg popular, re i(school) mle
Iteration 0:   log likelihood = -2556.3635
Iteration 1:   log likelihood = -2556.3612

Random-effects ML regression                    Number of obs      =      2000
Group variable (i): school                      Number of groups   =       100

Random effects u_i ~ Gaussian                   Obs per group: min =        16
                                                               avg =      20.0
                                                               max =        26

                                                Wald chi2(0)       =      0.00
Log likelihood  = -2556.3612                    Prob > chi2        =         .

------------------------------------------------------------------------------
     popular |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _cons |   5.307603   .0950194    55.86   0.000     5.121369    5.493838
-------------+----------------------------------------------------------------
    /sigma_u |   .9331052   .0684368                      .8081665    1.077359
    /sigma_e |   .7991726   .0129644                      .7741625    .8249907
         rho |   .5768565   .0367346                      .5039739    .6471936
------------------------------------------------------------------------------
Likelihood-ratio test of sigma_u=0: chibar2(01)= 1376.81 Prob>=chibar2 = 0.000
```

(Note: We would normally not recommend using gllamm for normally distributed responses since plenty of software exists for fitting such models without using approximation.)

The stata command is: **gllamm popular, i(school) adapt**

```
number of level 1 units = 2000
number of level 2 units = 100
Condition Number = 5.8576802
gllamm model
log likelihood = -2556.3612
------------------------------------------------------------------------------
     popular |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _cons |   5.307604   .0950217    55.86   0.000     5.121365    5.493843
------------------------------------------------------------------------------
Variance at level 1
-----------------------------------------------------------------------------
63867681 (.02072164)
Variances and covariances of random effects
-----------------------------------------------------------------------------
***level 2 (school)
    var(1): .87068762 (.12771943)
-----------------------------------------------------------------------------
```

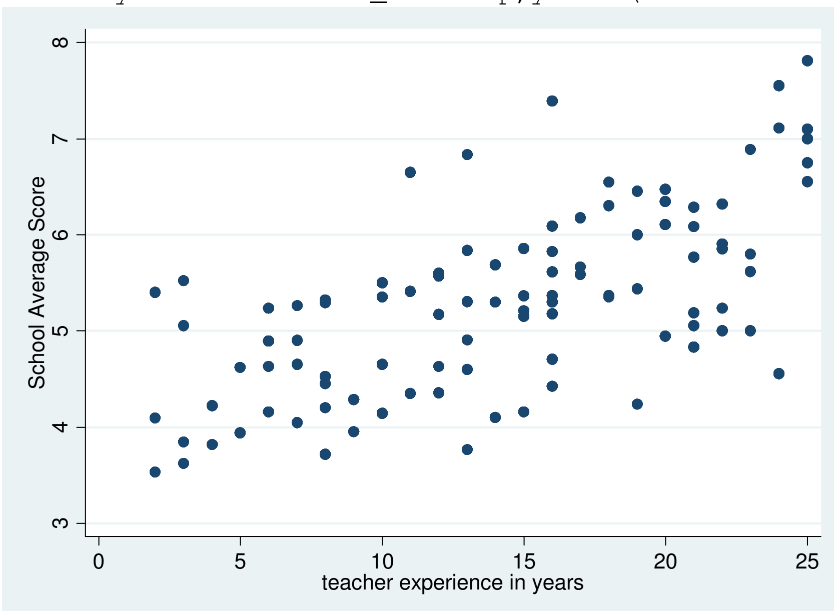
Q2. Do gender and teaching experience affect the self-rating score?

Gender is a 1st-stage covariate and teaching experience is a 2nd-stage covariate.

Exploratory analysis:

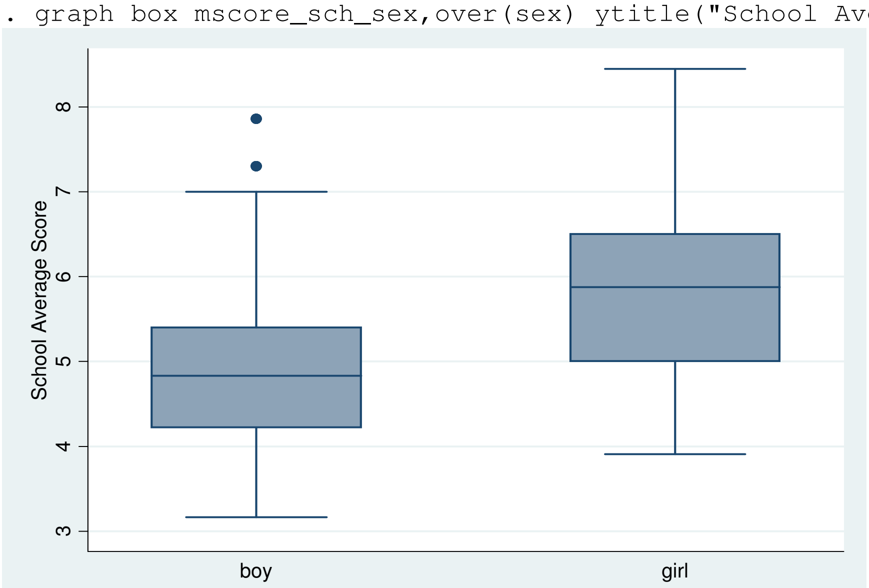
```
. sort school
. by school: egen mscore_sch=mean(popular)

. twoway scatter mscore_sch texp,ytitle("School Average Score")
```



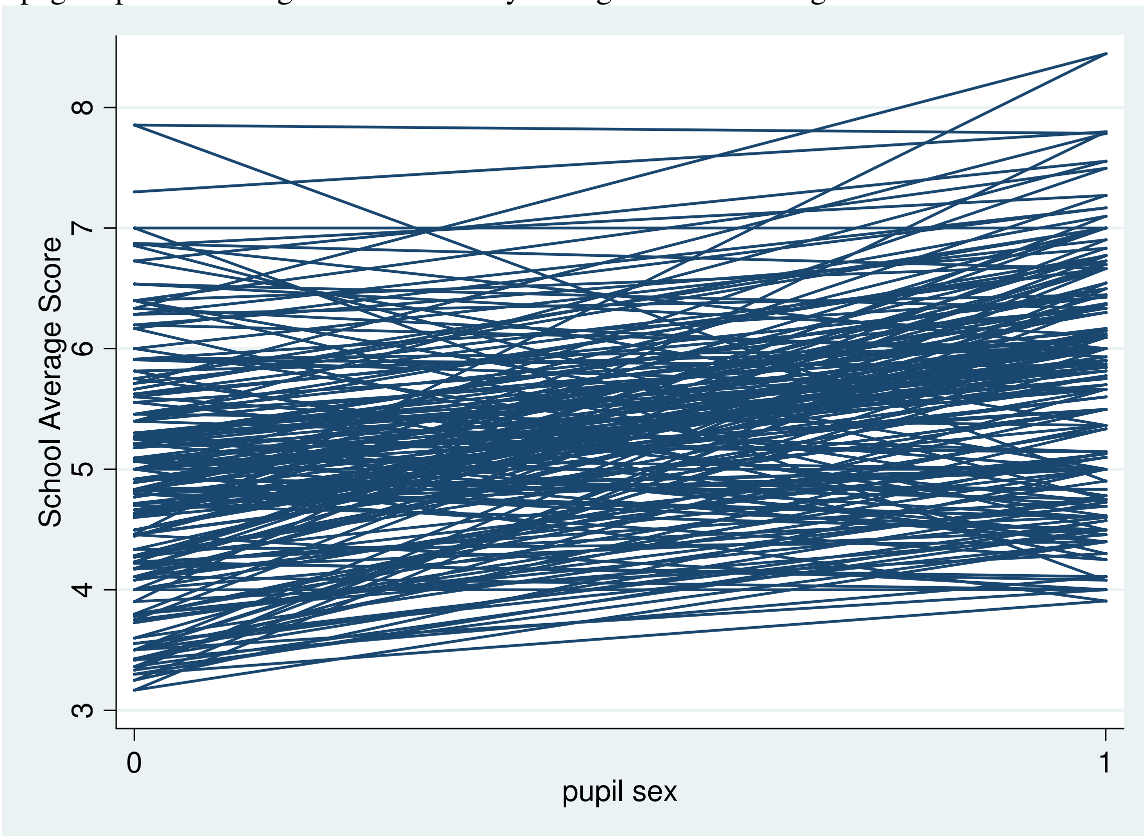
We find a positive association between teachers' experience and popularity score.

```
. sort school sex
. by school sex:egen mscore_sch_sex=mean(popular)
. sort sex
. graph box mscore_sch_sex,over(sex) ytitle("School Average Score")
```



We find girls give higher scores than boys.

Spagatti plot of average scores from boys and girls for each single school:



We find between-school heterogeneity of the gender effect on popularity score

Two-stage statistical model:

$$popular_{ij} = \beta_{0j} + \beta_{1j} Sex_{ij} + \varepsilon_{ij}$$

$$\beta_{0j} = \gamma_{00} + \gamma_{01} t\exp_j + \mu_{0j}$$

$$\beta_{1j} = \gamma_{10} + \mu_{1j}$$

$$\varepsilon_{ij} \sim N(0, \sigma^2)$$

$$\begin{pmatrix} \mu_{0j} \\ \mu_{1j} \end{pmatrix} \sim MVN(0, \Sigma)$$

Above equations can be written as a single complex regression model by substituting the equations for betas into the equation for the popularity.

$$popular_{ij} = \gamma_{00} + \gamma_{01} t\exp_j + \mu_{0j} + (\gamma_{10} + \mu_{1j}) \times Sex_{ij} + \varepsilon_{ij}\,,$$

where we model the popularity score as function of gender and teaching experience. We allow different baseline scores for different schools by using a random intercept, and we allow different gender effects for different schools by using a random slope for gender. We could fit random slope for teaching experience because it does not vary within school.

Rearrange above equation, we can see the fixed part is $\gamma_{00} + \gamma_{01} t\exp_{j} + \gamma_{10} \times Sex_{ij}$ since this segment contains the fixed coefficients.
Similarly, the random part is $\mu_{0j} + \mu_{1j} \times Sex_{ij} + \varepsilon_{ij}$. Since the covariate, sex, and the error term μ_{1j} is multiplied, the resulting total error will be different for different genders. This is a reason why analyzing the multi-level data with ordinary regression techniques does not work well.

The stata command used

```
. xtmixed popular texp sex || school: sex, cov(unstr) mle

Computing standard errors:

Mixed-effects ML regression                     Number of obs      =      2000
Group variable: school                          Number of groups   =       100

                                                Obs per group: min =        16
                                                               avg =      20.0
                                                               max =        26


                                                Wald chi2(2)       =    316.42
Log likelihood = -2130.5877                     Prob > chi2        =    0.0000

------------------------------------------------------------------------------
     popular |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        texp |   .1083526    .010112    10.72   0.000     .0885334    .1281718
         sex |   .8431752   .0593856    14.20   0.000     .7267815    .9595688
       _cons |   3.339973   .1591614    20.98   0.000     3.028022    3.651923
------------------------------------------------------------------------------

------------------------------------------------------------------------------
  Random-effects Parameters  |   Estimate   Std. Err.     [95% Conf. Interval]
-----------------------------+------------------------------------------------
school: Unstructured         |
                     sd(sex) |    .519327   .0483111      .4327695    .6231966
                   sd(_cons) |   .6344229   .0495562      .5443643    .7393807
             corr(sex,_cons) |   .0640675   .1309317     -.1911435    .3111648
-----------------------------+------------------------------------------------
                sd(Residual) |   .6264869   .0104455      .6063449     .647298
------------------------------------------------------------------------------
LR test vs. linear regression:       chi2(3) =  1274.41   Prob > chi2 = 0.0000

Note: LR test is conservative and provided only for reference
```

Fit this model using gllamm.

```
eq sch_s: sex
gen cons = 1
eq sch_c: cons
gllamm popular texp sex, i(school) adapt nrf(2) eq(sch_c sch_s)
```

```
number of level 1 units = 2000
number of level 2 units = 100
Condition Number = 40.498391
gllamm model
log likelihood = -2130.5659
------------------------------------------------------------------------------
     popular |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        texp |   .1084215   .0106385    10.19   0.000     .0875704    .1292726
         sex |   .8432452   .0587151    14.36   0.000     .7281656    .9583247
       _cons |   3.339099   .1651731    20.22   0.000     3.015366    3.662832
```

```
------------------------------------------------------------------------------
Variance at level 1
------------------------------------------------------------------------------
39241954 (.0130915)
Variances and covariances of random effects
------------------------------------------------------------------------------
***level 2 (school)
    var(1): .40328261 (.06227253)
    cov(1,2): .02171346 (.04195842) cor(1,2): .06576171
    var(2): .27033508 (.04961092)
------------------------------------------------------------------------------
```

It is important to always allow the random slope and random intercept to be correlated, otherwise, the fitted model will be biased.

Q3. Do teaching experience explains the between-school heterogeneity of gender effect?

Two-stage statistical model:

$$popular_{ij} = \beta_{0j} + \beta_{1j} Sex_{ij} + \varepsilon_{ij}$$

$$\beta_{0j} = \gamma_{00} + \gamma_{01} t\exp_j + \mu_{0j}$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11} t\exp_j + \mu_{1j}$$

$$\varepsilon_{ij} \sim N(0, \sigma^2)$$

$$\begin{pmatrix} \mu_{0j} \\ \mu_{1j} \end{pmatrix} \sim MVN(0, \Sigma)$$

Equivalently, the model could be written as

$$popular_{ij} = \gamma_{00} + \gamma_{01} t\exp_j + \mu_{0j} + (\gamma_{10} + \gamma_{11} t\exp_j + \mu_{1j}) \times Sex_{ij} + \varepsilon_{ij} .$$

Rearrange above equation, we can get the fixed part is $\gamma_{00} + \gamma_{01} t\exp_j + \gamma_{10} \times Sex_{ij} + \gamma_{11}(Sex_{ij} \times t\exp_j)$ since this segment contains the fixed coefficients and the random part is $\mu_{0j} + \mu_{1j} \times Sex_{ij} + \varepsilon_{ij}$.

Cross-level interaction of variable **sex** and **texp** is included. Notice this model takes very **LONG** time to run.

The STATA commands used are listed as followings:

```
gen gxt = sex*texp

xtmixed popular texp sex gxt || school: sex, cov(unstr) mle

Computing standard errors:

Mixed-effects ML regression                     Number of obs      =      2000
Group variable: school                          Number of groups   =       100
                                                Obs per group: min =        16
                                                               avg =      20.0
                                                               max =        26
                                                Wald chi2(3)       =    365.74
Log likelihood =  -2122.925                     Prob > chi2        =    0.0000
```

```
------------------------------------------------------------------------------
     popular |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        texp |   .1102293   .0101287    10.88   0.000     .0903774    .1300811
         sex |   1.329479   .1317029    10.09   0.000     1.071346    1.587612
         gxt |  -.0340251   .0083716    -4.06   0.000    -.0504331   -.0176172
       _cons |   3.313651   .1593869    20.79   0.000     3.001258    3.626044
------------------------------------------------------------------------------

------------------------------------------------------------------------------
  Random-effects Parameters  |   Estimate   Std. Err.     [95% Conf. Interval]
-----------------------------+------------------------------------------------
school: Unstructured         |
                    sd(sex) |   .4692521   .0458652      .3874439    .5683341
                  sd(_cons) |   .6347378   .0495438      .5446967    .7396631
            corr(sex,_cons) |   .0798403   .1247735     -.1645989    .3150401
-----------------------------+------------------------------------------------
                sd(Residual) |    .626432   .0104426      .6062956    .6472371
------------------------------------------------------------------------------
LR test vs. linear regression:       chi2(3) =  1269.28   Prob > chi2 = 0.0000

Note: LR test is conservative and provided only for reference
```

gllamm popular texp sex gxt, i(school) adapt nrf(2) eq(sch_c sch_s)

```
number of level 1 units = 2000
number of level 2 units = 100
Condition Number = 45.72526
gllamm model
log likelihood = -2122.9085
------------------------------------------------------------------------------
     popular |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        texp |   .1102169   .0099904    11.03   0.000      .090636    .1297977
         sex |    1.32949    .130912    10.16   0.000     1.072907    1.586073
         gxt |   -.034026   .0083388    -4.08   0.000    -.0503697   -.0176822
       _cons |   3.313841   .1566164    21.16   0.000     3.006879    3.620804
------------------------------------------------------------------------------
Variance at level 1
------------------------------------------------------------------------------
39236316 (.01308622)
Variances and covariances of random effects
------------------------------------------------------------------------------
***level 2 (school)
    var(1): .40543808 (.0627154)
    cov(1,2): .02386608 (.03657119) cor(1,2): .0795611
    var(2): .22194032 (.04304873)
------------------------------------------------------------------------------
```

The comparison of three models (Fitting using gllamm)

	Model 1	Model 2	Model 3
Fixed part	Estimate(SE)	Estimate(SE)	Estimate(SE)
Intercept	5.31(0.10)	3.34(0.16)	3.31(0.16)
Sex		0.84(0.06)	1.33(0.13)
texp		0.11(0.01)	0.11(0.01)
Texp*sex			-.03(0.01)
Random part			
σ_e^2	0.64(0.02)	0.39(0.01)	0.39(0.01)
σ_{u0}^2	0.87(0.13)	0.40(0.06)	0.41(0.06)
σ_{u1}^2		0.27(0.05)	0.22(0.04)

σ_{u01}		0.02(0.04)	0.02(0.04)

In this table, the intercept-only model (Model 1) estimates the intercept as 5.31, which is the average popularity across all schools and pupils. The variance of the pupils level residual errors, denoted by σ_e^2, is estimated as 0.64. The variance of the class level residual errors, denoted by σ_{u0}^2, is estimated as 0.87. The calculation of Z statistics for all parameter estimates shows that they are statistically significant at 0.05 level.

The second model includes pupil gender and teacher experience as predictors. The regression coefficients for both variables are significant. The coefficient for pupil gender is 0.84, this means that on average girls scores 0.84 points higher on the popularity measure. The coefficient for teacher experience is 0.11, which means for each year of the experience of the teachers, the average popularity score of the class goes up 0.11 points. The variance of the regression coefficient for pupil gender across classes is estimated as 0.27 with a standard error of 0.05. The covariance between the regression coefficients for gender and intercept is not significant.

The significant and quite large variance of the coefficient slope for pupil gender implies that the regression coefficient for pupil gender varies across the classes, and the value of 0.84 is just the expected value across all classes. The varying regression coefficients are assumed to follow a normal distribution. The variance of this distribution is estimated as 0.27.

The estimate of fixed coefficients for both model 2 and model 3 are similar except the regression slope for pupil gender, which is considerable larger in model 3. The interpretation remains same. The coefficient of the interaction between gender and teacher experience is estimated as -0.03, which is significant. The negative value means the difference between girls and boys is smaller with more experienced teachers. The variance component for pupil gender goes from 0.27 to 0.22, which means that model 3 explains some of the variation of the slope for gender pupil.