Introduction to LMER

Andrew Zieffler

Traditional Regression

- General form of the linear model

$$y_i = \beta_0 + \beta_1(x_{1i}) + \beta_2(x_{2i}) + \ldots + \beta_p(x_{pi}) + \epsilon_i$$

where

- y_i is the response for the i^{th} individual (i = 1, ..., N)
- x_{ki} is the k^{th} predictor (k = 1, ..., p)
- β_0 is the intercept
- β_k is the k^{th} regression coefficient (k > 0)
- ε_i is random error

Errors

- Errors are scatter around hyperplane defined by prediction equation
- For statistical inference errors assumed to be independent and normally distributed, with mean = 0 and constant variance, σ_{ϵ}^{2}.

Interpretation of Coefficients

- β_k indicates the change in the response for a unit increase in the k^{th} predictor holding all of the other predictors constant
- β_k indicates the strength of the relationship between the k^{th} predictor and the response controlling for all of other predictors constant

General Linear Model

- Various models subsumed based on nature of predictors
  - Use of dummy variables (ANOVA)
  - Mix of quantitative and dummy variables (ANCOVA)
  - Polynomials (Interaction models)

$$y_i = \beta_0 + \beta_1(x_{1i}) + \beta_2(x_{2i}) + \ldots + \beta_p(x_{pi}) + \epsilon_i$$

- Model presented is for a subject (note *i* subscript on *y*)
- Regression coefficients are at group-level
  - No *i* subscript
  - Index aggregate effects (do not vary across subjects)
  - Referred to as fixed-effects
- Includes subject-level and group-level terms and random error
- Focus typically on estimation and inference about fixed-effects

- Prediction equation

$$\hat{y}_i = \beta_0 + \beta_1(x_{1i}) + \beta_2(x_{2i}) + \ldots + \beta_p(x_{pi})$$

where

- $\hat{y}_i$ is the predicted or *fitted* response value

- Prediction is group-level enterprise in LM
- Traditional conceptualization
  - Predictor values set by analyst (or treated as such)
  - Assumption that multiple individuals share same values on predictor (in principle)
  - Fitted value for a subject is actually mean value for these subjects

Traditional Regression in R

- Examples only presented for illustration and to introduce syntax
- In anticipation of switching to LMER, only long-format of data is used
    - Because of lack of independence in long-data (multiple rows per subject) traditional regression would not be valid
    - Used here so comparison can be made later

Single Quantitative Predictor

- Use RStudio to import `Minneapolis-Long.csv` and assign it to `mpls.l`
- Use `grade` as single predictor of `read`
- Linear model is

$$y_i = \beta_0 + \beta_1(\text{grade}_i) + \epsilon_i$$

y is considered to be randomly sampled from population and grade is considered fixed.

- Use RStudio to import `Minneapolis-Long.csv` and assign it to `mpls.l`

- `lm()` function is used for regression
    - `response ~ 1 + predictor`
    - data frame
    - assign to object
- `summary()` function called on `lm` object

```
> lm.1 <- lm( read ~ 1 + grade, data = mpls.l )

> summary( lm.1 )
```

```
Call:
lm(formula = read ~ 1 + grade, data = mpls.l)

Residuals:
    Min      1Q  Median      3Q     Max 
-57.049  -7.512  -0.402  13.704  33.098 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)  183.915     13.238  13.893   <2e-16 ***
grade          4.427      2.056   2.153   0.0344 *  
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 19.53 on 78 degrees of freedom
  (8 observations deleted due to missingness)
Multiple R-squared: 0.05609,	Adjusted R-squared: 0.04399 
F-statistic: 4.635 on 1 and 78 DF,  p-value: 0.03442
```

```
> options( scipen = 999 )
```

turn off scientific notation

```
> library( ggplot2 )
```

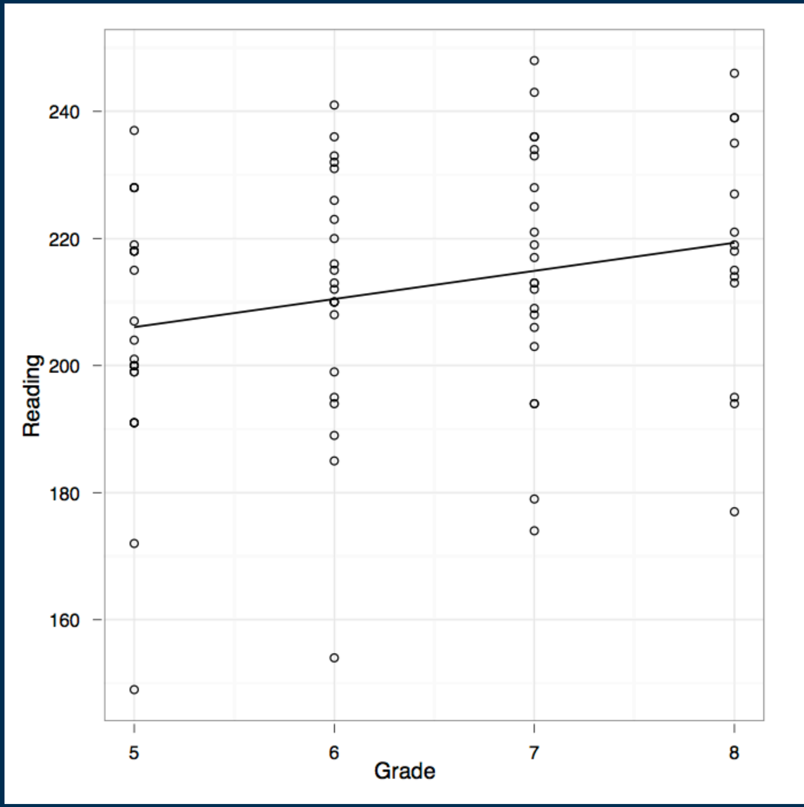
```
> mod.data <- fortify( lm.1 )
```

create data set from observations used to fit model

```
> mod.data
   read grade       .hat   .sigma        .cooksd  .fitted      .resid   .stdresid
1   172     5 0.03270510 19.25489 0.053129069473 206.0488 -34.0487805 -1.77277171
2   185     6 0.01385809 19.43627 0.012125655490 210.4756 -25.4756098 -1.31366788
3   179     7 0.01718404 19.21674 0.030064920968 214.9024 -35.9024390 -1.85446462
4   194     8 0.04268293 19.43225 0.039175926454 219.3293 -25.3292683 -1.32563952
5   200     5 0.03270510 19.64242 0.001676736177 206.0488  -6.0487805 -0.31493366
6   210     6 0.01385809 19.65484 0.000004226273 210.4756  -0.4756098 -0.02452515
7    209    7 0.01718404 19.64320 0.000812596932 214.9024  -5.9024390 -0.30487801
9    191    5 0.03270510 19.57742 0.010378430970 206.0488 -15.0487805 -0.78352446
10   199    6 0.01385809 19.61075 0.002460410614 210.4756 -11.4756098 -0.59174795
11   203    7 0.01718404 19.60724 0.003304335152 214.9024 -11.9024390 -0.61479534
```

Any row with NA dropped from analysis

```
> ggplot( data = mod.data, aes( x = grade, y = read ) ) +
    geom_point( shape = 1 ) +
    geom_line( aes( x = grade, y = .fitted ) ) +
    theme_bw() +
    scale_x_continuous( name = "Grade", breaks = 5:8 ) +
    scale_y_continuous( name = "Reading" )
```



- Consistent with slope coefficient, line has positive slope
- Predicted mean reading achievement score increases over grade

ANCOVA as Regression

- Goal is to determine if there are differences in intercepts between groups
- Quantitative predictor and categorical predictor in model
- Dummy coded categorical predictor(s)

```
> mpls.l$dadv <- ifelse( mpls.l$risk == "ADV", 0, 1 )
```

logical expression

value if true

value if false

```
> head( mpls.l )

  subid risk gen eth ell sped  att grade read dadv
1     1  HHM   F Afr   0    N 0.94     5  172    1
2     1  HHM   F Afr   0    N 0.94     6  185    1
3     1  HHM   F Afr   0    N 0.94     7  179    1
4     1  HHM   F Afr   0    N 0.94     8  194    1
5     2  HHM   F Afr   0    N 0.91     5  200    1
6     2  HHM   F Afr   0    N 0.91     6  210    1
```

0 = advantaged

1 = disadvantaged

- Use `grade` and `dadv` as predictors of `read`
- Linear model is

$$y_i = \beta_0 + \beta_1(\mathrm{grade}_i) + \beta_2(\mathrm{dadv}_i) + \epsilon_i$$

```
> lm.2 <- lm( read ~ 1 + grade + dadv, data = mpls.l )

> summary( lm.2 )
```

```
Call:
lm(formula = read ~ 1 + grade + dadv, data = mpls.l)

Residuals:
    Min      1Q  Median      3Q     Max 
-48.035  -9.817   1.548   8.982  41.850 

Coefficients:
            Estimate Std. Error t value             Pr(>|t|)    
(Intercept)  193.886     11.649  16.644 < 0.0000000000000002 ***
grade          4.557      1.784   2.554               0.0126 *  
dadv         -19.638      3.809  -5.156            0.0000019 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 16.95 on 77 degrees of freedom
  (8 observations deleted due to missingness)
Multiple R-squared: 0.2983,	Adjusted R-squared: 0.2801 
F-statistic: 16.37 on 2 and 77 DF,  p-value: 0.000001191
```

$$\hat{y}_i = \beta_0 + \beta_1(\text{grade}_i) + \beta_2(\text{dadv}_i)$$

- Conditioning on dadv

$$\hat{y}_i = \beta_0 + \beta_1(\text{grade}_i)$$ if dadv = 0

$$\hat{y}_i = (\beta_0 + \beta_2) + \beta_1(\text{grade}_i)$$ if dadv = 1

- Substitute in coefficients

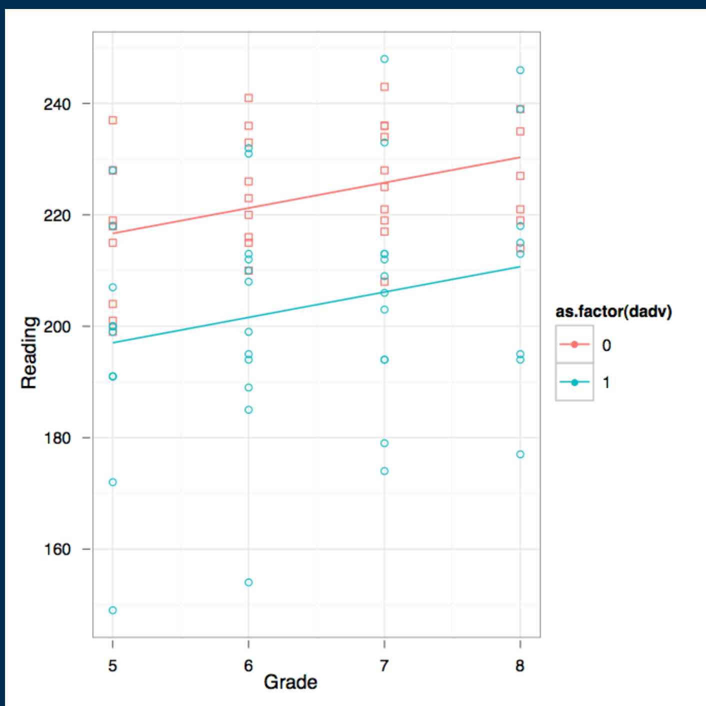
$$\hat{\beta}_0 = 193.89 \qquad \hat{\beta}_1 = 4.56 \qquad \hat{\beta}_2 = -19.64$$

$$\hat{y}_i = 193.89 + 4.56(\text{grade}_i)$$ if dadv = 0

$$\hat{y}_i = 174.25 + 4.56(\text{grade}_i)$$ if dadv = 1

```
> mod.data2 <- fortify( lm.2 )

> ggplot( data = mod.data2, aes( x = grade, y = read,
          group = dadv, color = as.factor( dadv ) ) ) +
    geom_point( aes( shape = dadv ) ) +
    geom_line( aes( x = grade, y = .fitted ) ) +
    theme_bw() +
    scale_x_continuous( name = "Grade", breaks = 5:8 ) +
    scale_y_continuous( name = "Reading" )
```



- Difference in intercepts, after controlling for grade
- In LMER, when static predictor is in model along with time predictor, called *intercept effect*

Interaction Model

- Goal is to determine if there are differences in intercepts *and* slopes between groups
- Quantitative predictor and categorical predictor (dummy coded) and interaction term in model

- Use grade and dadv and grade*dadv as predictors of read
- Linear model is

$$y_i = \beta_0 + \beta_1(\text{grade}_i) + \beta_2(\text{dadv}_i) + \beta_3(\text{grade}_i)(\text{dadv}_i) + \epsilon_i$$

```
> lm.3 <- lm( read ~ 1 + grade + dadv + grade:dadv, data = mpls.l )
> summary( lm.3 )
```

```
Call:
lm(formula = read ~ 1 + grade + dadv + grade:dadv, data = mpls.l)

Residuals:
    Min      1Q  Median      3Q     Max 
-47.264  -8.390   0.391   9.817  41.275 

Coefficients:
            Estimate Std. Error t value            Pr(>|t|)    
(Intercept)  201.058     17.280  11.636 <0.0000000000000002 ***
grade          3.425      2.691   1.273               0.207    
dadv         -32.555     23.216  -1.402               0.165    
grade:dadv     2.035      3.608   0.564               0.574    
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 17.02 on 76 degrees of freedom
  (8 observations deleted due to missingness)
Multiple R-squared: 0.3013,	Adjusted R-squared: 0.2737 
F-statistic: 10.92 on 3 and 76 DF,  p-value: 0.00000481
```

$$\hat{y}_i = \beta_0 + \beta_1(\text{grade}_i) + \beta_2(\text{dadv}_i) + \beta_3(\text{grade}_i)(\text{dadv}_i)$$

- Conditioning on dadv

$$\hat{y}_i = \beta_0 + \beta_1(\text{grade}_i)$$ if dadv = 0

$$\hat{y}_i = (\beta_0 + \beta_2) + (\beta_1 + \beta_3)(\text{grade}_i)$$ if dadv = 1

- Substitute in coefficients

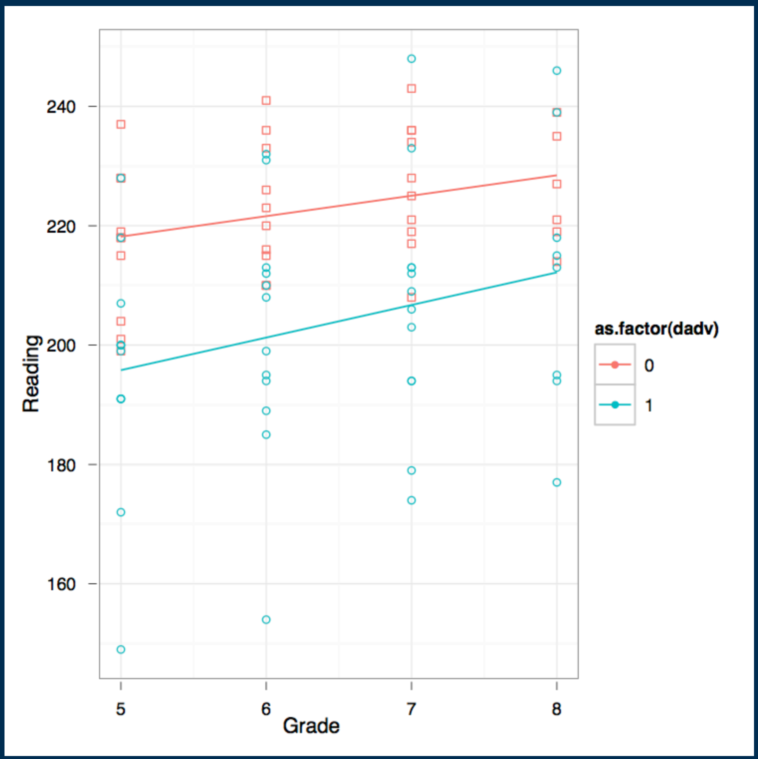
$$\hat{\beta}_0 = 201.06 \qquad \hat{\beta}_1 = 3.43 \qquad \hat{\beta}_2 = -32.56 \qquad \hat{\beta}_3 = 2.04$$

$$\hat{y}_i = 201.06 + 3.43(\text{grade}_i)$$ if dadv = 0

$$\hat{y}_i = 168.50 + 5.46(\text{grade}_i)$$ if dadv = 1

```
> mod.data3 <- fortify( lm.3 )

> ggplot( data = mod.data3, aes( x = grade, y = read,
         group = dadv, color = as.factor( dadv ) ) ) +
    geom_point( aes( shape = dadv ) ) +
    geom_line( aes( x = grade, y = .fitted ) ) +
    theme_bw() +
    scale_x_continuous( name = "Grade", breaks = 5:8 ) +
    scale_y_continuous( name = "Reading" )
```



- Difference in intercepts, but not slopes, after controlling for grade
- In LMER called *intercept and slope effects*

Linear Mixed Effects Regression

- Review of models fitted
  - Unconditional model: `grade` as predictor of `read`
  - Intercept effect: `grade` and `dadv` as predictors of `read`
  - Intercept and slope effects: `grade`, `dadv`, and `grade*dadv` as predictors of `read`
- These fundamental models address typical research questions

Research Questions

- Unconditional model
  - Focus is change over time, ignoring any other predictors
- Intercept effects/slope effects
  - Focus is change over time, conditional on level of some predictor(s)
  - Is starting point the same? Is change curve the same?

Non-Independence

- Data in long format has repeated rows per subject
  - Independence assumes each row is different subject
- Problem: How are subjects properly associated with their repeated measures?
  - Solution: Extend regression model to include *random effects*

Random Effects

- Random effects embed subject-specific model within the larger regression model
    - Subject-specific model defined for each block of repeated observations
    - Based on a block ID (subject ID) same for all repeated measures that belong to common subject (e.g., `subid`)
- After blocks of subjects are accounted for, random effects allow subjects to be treated as independent

Extension of Linear Model

- Terms appearing in both LM and LMER have similar interpretations
- Both models have fixed effects and random error
  - LMER adds random effects
  - Models that include both fixed effects *and* random effects referred to as *mixed effects models*

Notation of LMER

- Expanded notation
  - Subject index: $i = 1, ..., N$
  - Time point index: $j = 1, ..., n_i$
  - Subject index used within time point index to denote missing data
    - Subject 1 (no missing data): $n_1 = 4$
    - Subject 2 (missing grade 8): $n_2 = 3$

- Consider analog of unconditional model (use `grade` as predictor of `read`)
- LMER model is

$$y_{ij} = (\beta_0 + b_{0i}) + (\beta_1 + b_{1i})(\mathrm{grade}_{ij}) + \epsilon_{ij}$$

where

  - b_{0i} and b_{1i} are random effects
  - β_0 and β_1 are fixed effects
  - ε_{ij} is random error

- Not compulsory to include both random effects
  - If higher order random effects are included, all lower order random effects must also be included
- All random effects have a subject subscript and no time subscript
  - Summarize across repeated measures for each subject
- Fixed effects have same interpretations
  - β_0 is group-level intercept–predicted level at grade=0
  - β_1 is group-level slope–predicted change in mean reading score for each grade increase

- Estimation and inference of fixed effects is of interest to applied researchers
  - Tertiary roles: Variance of errors and estimation/inference of random effects
- Estimation typically does not use OLS
  - Maximum likelihood methods commonly employed
  - Different estimation methods produce different fixed effects estimates with missing data
  - Larger sample sizes produce less of a difference in these estimates
  - Estimation of SEs do not have same large-sample correspondence
  - Additional parameters in LMER (i.e., random effects) influence SE

LMER as Multilevel Model

- Multilevel expression of LMER model explicitly separates the between-subject and within-subject variation
  - Level 1 model: Within-subjects aspects of model
  - Level 2 model: Between-subjects aspects of model

Level 1: $$y_{ij} = \beta^*_{0i} + \beta^*_{1i}(\text{grade}_{ij}) + \epsilon_{ij}$$

Level 2: $$\left\{ \begin{array}{l} \beta^*_{0i} = \beta_0 + b_{0i} \\ \beta^*_{1i} = \beta_1 + b_{1i} \end{array} \right.$$

- Level 1 model is subject-specific change curve
  - β_{0i}^* is the intercept for the *i*th subject
  - β_{1i}^* is the slope for the *i*th subject
  - ϵ_{ij} are the random errors around the *i*th subject's regression line
- Only source of variation in Level 1 model is within-subject variation (pertaining to repeated measures)
  - Time predictors and dynamic covariates appear exclusively in Level 1 model
  - Any variation that is within-subject not accounted for by time predictors and dynamic covariates absorbed into ϵ_{ij}

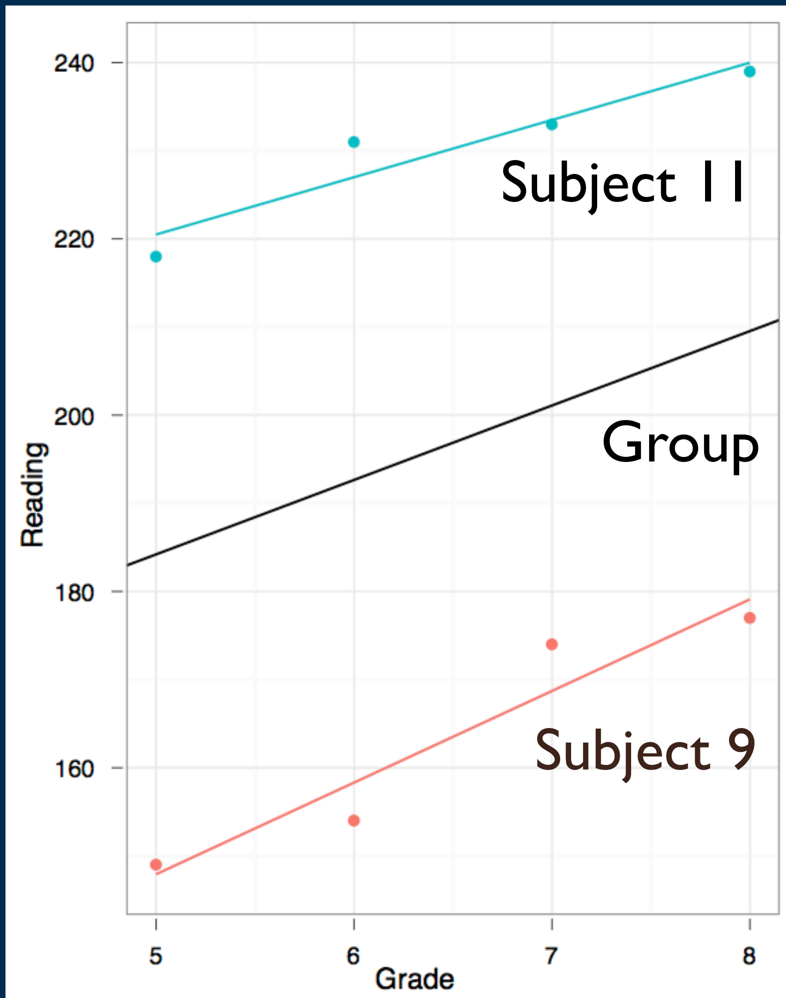
- Level 2 model are group-level equations involving fixed effects
  - Response variables are the Level 1 intercepts and slopes
  - Number of Level 2 equations determined by number of parameters in the Level 1 model
  - Static covariates that account for between-subject variation exclusively appear in Level 2 model
- Error terms in Level 2 model are the random effects
  - Indifferent to the number of fixed effects at Level 2
  - Can only be as many random effects as there are Level 2 equations
  - Not every Level 2 equation needs to have random effect

- Substitute Level 2 equations into Level 1 equation to get LMER model
  - To develop LMER model, often helpful to begin with multilevel model
  - Especially true when subject-specific change curves are non-linear or there are many dynamic covariates
  - LMER model maps to syntax used in `lmer()` function

- Re-expression of multilevel model helps clarify nature of random effects

$$b_{0i} = \beta^*_{0i} - \beta_0$$
$$b_{1i} = \beta^*_{1i} - \beta_1$$

- Random effect is individual deviation from group-level fixed effect
  - b_{0i} is the discrepancy between the individual and group-level intercept for the *i*th subject
  - b_{1i} is the discrepancy between the individual and group-level slope for the *i*th subject
  - Random effects can be negative, positive, or 0



Intercepts

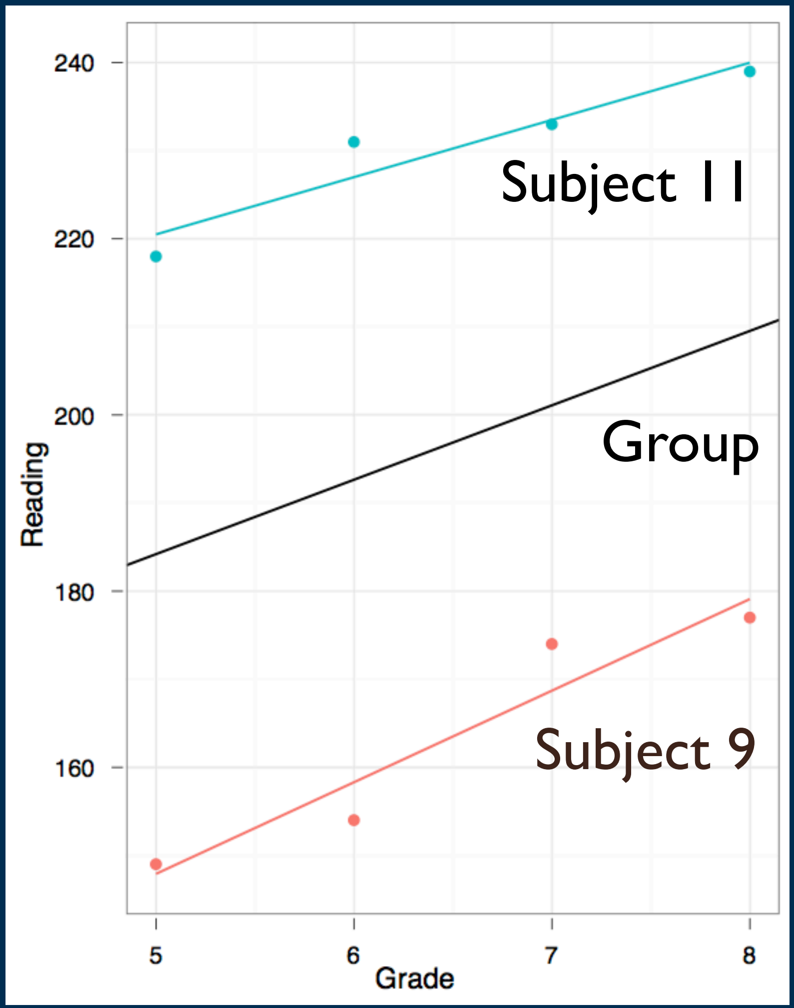
Subject 11: $\beta^*_{011} = 220.5$

$b_{011} = 36.3$

Group: $\beta_0 = 184.2$

$b_{09} = -36.3$

Subject 9: $\beta^*_{09} = 147.9$



Slopes

Subject 11: $\beta^*_{111} = 6.5$

$b_{111} = -1.95$

Group: $\beta_1 = 8.45$

$b_{19} = 1.95$

Subject 9: $\beta^*_{19} = 10.4$

Two Types of Random Effects

- In multilevel model, the β^*_{ki} coefficients are referred to as *mean-uncorrected* random effects, and the b_{ki} coefficents are referred to as *mean-corrected* random effects

$$E(\beta^*_{ki}) = \beta_k$$

mean is a fixed effect

$$\begin{aligned} E(b_{ki}) &= E(\beta^*_{ki} - \beta_k) \\ &= E(\beta^*_{ki}) - E(\beta_k) \\ &= \beta_k - \beta_k \\ &= 0 \end{aligned}$$

mean is 0

- Distinction between mean-corrected and mean-uncorrected random effects is not important in many LMER analyses
  - Variances and covariances are focus, rather than actual values of random effects
  - Adding a constant does not affect computation of variances, covariances or correlations
- Distinction is important in estimating change curve for a subject
  - Focus is decidedly on β^*_{ki} rather than b_{ki}

Random Effects as Errors

- Random effects can be conceived of as error terms (see multilevel model)
- Can also be emphasized in LMER model

$$
\begin{aligned}
y_{ij} &= (\beta_0 + b_{0i}) + (\beta_1 + b_{1i})(\text{grade}_{ij}) + \epsilon_{ij} \\
&= \beta_0 + b_{0i} + \beta_1(\text{grade}_{ij}) + b_{1i}(\text{grade}_{ij}) + \epsilon_{ij} \\
&= \beta_0 + \beta_1(\text{grade}_{ij}) + \boxed{(b_{0i} + b_{1i}(\text{grade}_{ij}) + \epsilon_{ij})}
\end{aligned}
$$

collect random effects with random error

$$y_{ij} = \beta_0 + \beta_1(\text{grade}_{ij}) + (b_{0i} + b_{1i}(\text{grade}_{ij}) + \epsilon_{ij})$$

- This expression of the LMER is representation of syntax used in `lmer()`
  - Requires separation of fixed effects and random effects
- One step further

$$y_{ij} = \beta_0 + \beta_1(\text{grade}_{ij}) + \epsilon^*_{ij}$$

- Very similar to LM
  - Error term in LM reflects only between-subjects deviations
  - Error term in LMER reflects between-subjects and within-subjects deviations

- Additionally for fitted equation, random effects drop out with random error

$$\begin{aligned}\hat{y}_{ij} &= E(y_{ij}) \\ &= E(\beta_0) + E(\beta_1(\text{grade}_{ij})) + E(\epsilon^*_{ij}) \\ &= \beta_0 + \beta_1(\text{grade}_{ij})\end{aligned}$$

- This follows from the assumptions put on the random effects and random error

Assumptions

- Statistical inference for LMER predicated ϵ^*_{ij} on assumptions about the error term,
- More complex than LM since error term is composed of three different components

Assumptions in EPsy 8282

- Random effects have a joint normal distribution
- Random errors are normally distributed
- Random errors are independent between time points, and have constant variance over time
- Random effects are correlated, but independent of the random errors
- Random effects and random errors each have mean = 0

- Random error has constant variance σ_ϵ^2
  - σ_ϵ^2 is variance associated with subject-level change curve
  - Magnitude is influenced by time predictors and dynamic covariates in model
  - Adding these predictors generally will reduce value
- Random effects have variances $Var(b_{0i})$ and $Var(b_{1i})$
  - Index subject variability in intercepts and slopes
  - Magnitudes influenced by static covariates included in model
  - Adding static covariates generally reduces values
  - If these are both 0, using LMER will be pointless

- Covariance between random effects $Cov(b_{0i}, b_{1i})$
  - Covariance transformed to correlation by dividing by product of square root of variances

$$Corr(b_{0i}, b_{1i}) = \frac{Cov(b_{0i}, b_{1i})}{\sqrt{Var(b_{0i})}\sqrt{Var(b_{1i})}}$$

  - Correlation indexes magnitude of relationship between subject intercepts and slopes
    - Positive correlation indicates that larger intercepts tend to be associated with larger slopes
    - Negative correlation indicates that larger intercepts tend to be associated with smaller slopes

Random Effects and Correlated Observations

- Random effects index individual deviations from the group-level fixed effects
- How do they provide a model for the dependency found in repeated measures?

- Consider variance-covariance matrix among repeated measures of response variable
- Consists of variances and covariances among reading scores at the four grades

$$
\begin{array}{cc}
\text{Grade} & \begin{array}{cccc} \quad 5 \quad & \quad 6 \quad & \quad 7 \quad & \quad 8 \quad \end{array} \\
\begin{array}{c} 5 \\ 6 \\ 7 \\ 8 \end{array} &
\begin{pmatrix}
Var(y_{i1}) & Cov(y_{i1}, y_{i2}) & Cov(y_{i1}, y_{i3}) & Cov(y_{i1}, y_{i4}) \\
Cov(y_{i2}, y_{i1}) & Var(y_{i2}) & Cov(y_{i2}, y_{i3}) & Cov(y_{i2}, y_{i4}) \\
Cov(y_{i3}, y_{i1}) & Cov(y_{i3}, y_{i2}) & Var(y_{i3}) & Cov(y_{i3}, y_{i4}) \\
Cov(y_{i4}, y_{i1}) & Cov(y_{i4}, y_{i2}) & Cov(y_{i4}, y_{i3}) & Var(y_{i14})
\end{pmatrix}
\end{array}
$$

- When covariances equal 0, there is no dependency between time points
  - This would be strange for longitudinal data
  - We typically assume covariances that are non-zero
- LMER models account for non-zero covariances by way of random effects and random error
  - It can be shown that variances and covariances among repeated measures is a function of

    $$Var(b_{0i}) \quad Var(b_{1i}) \quad Cov(b_{0i}, b_{1i}) \quad \sigma_\epsilon^2$$

  - Thus random effects provide a model for the correlated observations characterizing longitudinal data

Estimating LMER Model

- LMER models estimated using the `lmer()` function from the **lme4** package
- Uses syntax similar to the `lm()` function
    - Fixed effects uses exact same syntax
    - Random effects appear in parentheses and reference subject ID variable

Time as Single Predictor

$$y_{ij} = \beta_0(1) + \beta_1(\text{grade}_{ij}) + (b_{0i}(1) + b_{1i}(\text{grade}_{ij}) + \epsilon_{ij})$$

```
> lmer( read ~ 1 + grade + ( 1 + grade | subid ), data = mpls.l,
        REML = FALSE)
```

Subject ID

```
> library( lme4 )

> lmer.1 <- lmer( read ~ 1 + grade + ( 1 + grade | subid ),
                  data = mpls.l, REML = FALSE )


> summary( lmer.1 )
```

```
Linear mixed model fit by maximum likelihood
Formula: read ~ 1 + grade + (1 + grade | subid)
   Data: mpls.l
   AIC BIC logLik deviance REMLdev
 583.7 598 -285.8    571.7   565.8
Random effects:
 Groups   Name        Variance Std.Dev. Corr
 subid    (Intercept) 740.4670 27.2115
          grade         6.9662  2.6394  -0.744
 Residual              18.3152  4.2796
Number of obs: 80, groups: subid, 22

Fixed effects:
            Estimate Std. Error t value
(Intercept) 181.3335     6.5614  27.636
grade         4.8823     0.7417   6.582

Correlation of Fixed Effects:
      (Intr)
grade -0.799
```

```
> round(summary(lm.1)$coeff, 4)
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 183.9146    13.2382 13.8927   0.0000
grade         4.4268     2.0562  2.1529   0.0344

> round(summary(lmer.1)@coefs, 4)
            Estimate Std. Error t value
(Intercept) 181.3335     6.5614 27.6363
grade         4.8823     0.7417  6.5822
```

- Fixed effects estimates are similar
- SEs for LMER model 1/2 the size of LM SEs
- This results in larger *t*-values for LMER model
- Researcher would definitely want to use LMER rather than LM when there is dependency

No time predictor (random intercepts model)

```
> lmer.0 <- lmer(read ~ 1 + (1 | subid), data = mpls.l, REML = FALSE)

> round(summary(lmer.0)@sigma ^ 2, 2)
[1] 66.2
```

Time predictor

```
> round(summary(lmer.1)@sigma ^ 2, 2)
[1] 18.32
```

- Error variance is influenced by time predictors
- Since error variance is reduced after including grade in the model, it suggests grade accounts for a portion of the within-subjects variation
- Without grade in the model, the unaccounted for variation is absorbed into the residuals

Anchoring the Intercept

- Using grade as time predictor produces interpretation of β_0 as estimated group reading level for the 0^{th} grade
- To facilitate more meaningful interpretations, linear transformation is performed on time predictor
- Often convenient to anchor intercept to first time point

To anchor intercept to first time point (grade 5), minimum value of time predictor must equal 0

```
> mpls.l$grade5 <- mpls.l$grade - 5

> lmer.1a <- lmer(read ~ grade5 + (1 + grade5 | subid), mpls.l, REML = FALSE)
> summary( lmer.1a )
```

```
Linear mixed model fit by maximum likelihood
Formula: read ~ grade5 + (1 + grade5 | subid)
   Data: mpls.l
   AIC BIC logLik deviance REMLdev
 583.7 598 -285.8    571.7   565.8
Random effects:
 Groups   Name        Variance Std.Dev. Corr
 subid    (Intercept) 380.5857 19.5086
          grade5        6.9662  2.6394  -0.361
 Residual              18.3153  4.2796
Number of obs: 80, groups: subid, 22

Fixed effects:
            Estimate Std. Error t value
(Intercept) 205.7451     4.2322   48.61
grade5        4.8823     0.7417    6.58

Correlation of Fixed Effects:
       (Intr)
grade5 -0.363
```

β_0 is now estimated group reading level for the 5th grade

To anchor intercept to last time point (grade 8), maximum value of time predictor must equal 0

```
> lmer.1b <- lmer(read ~ I( grade - 8 ) + (1 + I( grade - 8 ) | subid), mpls.l, REML = FALSE)
> summary( lmer.1b )

Linear mixed model fit by maximum likelihood
Formula: read ~ I(grade - 8) + (1 + I(grade - 8) | subid)
   Data: mpls.l
   AIC BIC logLik deviance REMLdev
 583.7 598 -285.8    571.7   565.8
Random effects:
 Groups   Name         Variance Std.Dev. Corr
 subid    (Intercept)  331.8442 18.2166
          I(grade - 8)   6.9662  2.6394  0.048
 Residual               18.3153  4.2796
Number of obs: 80, groups: subid, 22

Fixed effects:
             Estimate Std. Error t value
(Intercept)  220.3921     4.0039   55.04
I(grade - 8)   4.8823     0.7417    6.58

Correlation of Fixed Effects:
           (Intr)
I(grade-8) 0.172
```

β_0 is now estimated group reading level for the 8th grade

```
                 grade     grade5     grade8
(Intercept) 181.333540 205.745124 220.392076
grade         4.882317   4.882314   4.882318
```

- Nothing sacrosanct about intercept, based on interpretation
- Transformation must occur in both fixed and random effects
- Results show larger intercept as increase in grade where anchoring occurs
- Slight changes in slopes (due to differences in variances and covariances of random effects in the models)
- Interpretation of $Var(b_{0i})$ also changes–variance of intercept at X grade
- Since time transformation is a re-scaling, often desirable to keep random intercept term in model, even if estimated variance is 0 (consider fan-shaped variance)

LMER with Static Covariates

- Consider research where goal is to examine effect of risk (`dadv`) *on intercepts*
- Similar to ANCOVA model, `dadv` is included in model as effect

$$y_{ij} = \beta_0 + \beta_1(\text{grade}_{ij}) + \beta_2(\text{dadv}_i) + (b_{0i} + b_{1i}(\text{grade}_{ij}) + \epsilon_{ij})$$

- No random effect is introduced for β_2
- Remember, random effects are only specified for parameters in the Level 1 model (time predictors or dynamic covariates–not static covariates)

Level 1:

$$y_{ij} = \beta^*_{i0} + \beta^*_{i1}(\text{grade}_{ij}) + \epsilon_{ij})$$

Level 2:

$$\begin{cases} \beta^*_{i0} = \beta_0 + \beta_2(\text{dadv}_i) + b_{0i} \\ \beta^*_{i1} = \beta_1 + b_{1i} \end{cases}$$

- Risk predictor appears in Level 2 model since it predicts between-subject variation in intercepts
- β_2 is group-level effect (mean difference since it is dummy coded) of being disadvantaged on the intercept

Include static predictor of risk

```
> lmer.2 <- lmer( read ~ grade + dadv + ( 1 + grade | subid ), mpls.l, REML = FALSE)
> summary( lmer.2 )
```

```
Formula: read ~ grade + dadv + (1 + grade | subid)
   Data: mpls.l
   AIC   BIC logLik deviance REMLdev
 577.9 594.6   -282    563.9   552.7
Random effects:
 Groups   Name        Variance Std.Dev. Corr
 subid    (Intercept) 600.3700 24.5024
          grade         7.1643  2.6766  -0.782
 Residual              18.1295  4.2579
Number of obs: 80, groups: subid, 22

Fixed effects:
            Estimate Std. Error t value
(Intercept) 192.4531     7.0479  27.306
grade         4.8836     0.7466   6.541
dadv        -20.3988     6.6515  -3.067

Correlation of Fixed Effects:
      (Intr) grade
grade -0.717
dadv  -0.513 -0.003
```

- Now consider research where goal is to examine effect of risk (`dadv`) *on both intercepts and slopes*
- Similar to interaction model, `dadv` and `dadv:grade` are included in model as effects

Level 1: $$y_{ij} = \beta^*_{i0} + \beta^*_{i1}(\text{grade}_{ij}) + \epsilon_{ij}$$

Level 2: $$\begin{cases} \beta^*_{i0} = \beta_0 + \beta_2(\text{dadv}_i) + b_{0i} \\ \beta^*_{i1} = \beta_1 + \beta_3(\text{dadv}_i) + b_{1i} \end{cases}$$

- β_3 indexes relationship between subjects' slopes and risk predictor

- Substituting

$$
\begin{aligned}
y_{ij} &= [\beta_0 + \beta_2(\mathrm{dadv}_i) + b_{0i}] + [\beta_1 + \beta_3(\mathrm{dadv}_i) + b_{1i}](\mathrm{grade}_{ij}) + \epsilon_{ij} \\
&= \beta_0 + \beta_2(\mathrm{dadv}_i) + b_{0i} + \beta_1(\mathrm{grade}_{ij}) + \beta_3(\mathrm{dadv}_i)(\mathrm{grade}_{ij}) + b_{1i}(\mathrm{grade}_{ij}) + \epsilon_{ij} \\
&= \beta_0 + \beta_1(\mathrm{grade}_{ij}) + \beta_2(\mathrm{dadv}_i) + \beta_3(\mathrm{dadv}_i)(\mathrm{grade}_{ij}) + (b_{0i} + b_{1i}(\mathrm{grade}_{ij}) + \epsilon_{ij})
\end{aligned}
$$

```
> lmer.3 <- lmer( read ~ grade + dadv + grade:dadv + ( 1 + grade ) | subid ),
                 mpls.l, REML = FALSE)
> summary( lmer.3 )
```

```
Linear mixed model fit by maximum likelihood
Formula: read ~ grade + dadv + grade:dadv + (1 + grade | subid)
   Data: mpls.l
   AIC   BIC logLik deviance REMLdev
 579.8 598.8 -281.9    563.8   549.9
Random effects:
 Groups   Name        Variance Std.Dev. Corr
 subid    (Intercept) 593.6171 24.3643
          grade         6.9489  2.6361  -0.779
 Residual              18.2444  4.2713
Number of obs: 80, groups: subid, 22

Fixed effects:
            Estimate Std. Error t value
(Intercept)  194.571      8.973  21.685
grade          4.570      1.106   4.130
dadv         -24.258     12.115  -2.002
grade:dadv     0.571      1.489   0.383

Correlation of Fixed Effects:
           (Intr) grade  dadv
grade      -0.837
dadv       -0.741  0.620
grade:dadv  0.622 -0.743 -0.836
```

Initial Status as Static Covariate

- Can statistically control for initial status
- Useful when first measurement is baseline in non-experimental studies
  - For example, use reading in 5th grade as static covariate
  - Examine change from grades 6 to 8

- Need to extract 5th grade reading scores from `mpls.l` data set

```
> grade5 <- subset( mpls.l, grade == 5, select = c( subid, read ) )
> names( grade5 )[2] <- "read.int"
```

- Need to extract rows for 6th–8th grade from `mpls.l` data set

```
> grade6to8 <- subset( mpls.l, grade != 5 )
```

- Merge two new data frames together

```
> mpls.l.2 <- merge( grade5, grade6to8, by = "subid" )
```

```
> head(mpls.l.2)
  subid read.int risk gen eth ell sped  att grade read dadv grade5
1     1      172  HHM   F Afr   0    N 0.94     6  185    1      1
2     1      172  HHM   F Afr   0    N 0.94     7  179    1      2
3     1      172  HHM   F Afr   0    N 0.94     8  194    1      3
4     2      200  HHM   F Afr   0    N 0.91     6  210    1      1
5     2      200  HHM   F Afr   0    N 0.91     7  209    1      2
6     2      200  HHM   F Afr   0    N 0.91     8   NA    1      3
```