Applied longitudinal Data Analysis Workshop 1: Multilevel Model for Change

Tyler H. Matta

Educational Methodology, Policy & Leadership
Center for Assessment Statistics & Evaluation

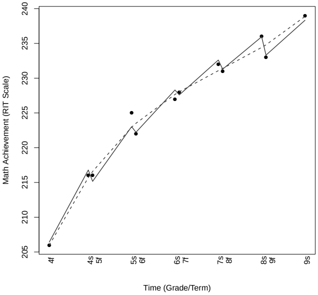
October 4, 2014

Workshop Overview

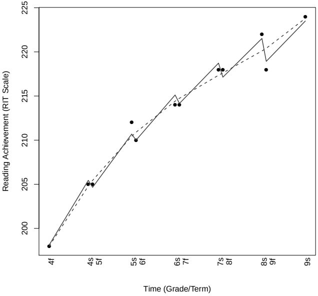
- Longitudinal data and the multilevel model for change
- Strategies for model fitting
- Time-varying covariates
- Non-linear growth

Personal Experience with Multilevel Models for Change

$$y_g = \underbrace{x_g^F \beta^F}_{\text{fall status}} + \underbrace{x_g^{FS} \beta^{FS}}_{\text{fall-spring gain}} + \nu_g$$



(a) Math Growth



(b) Reading Growth

Figure : "Additive Polynomial" Models (Thum & Matta (n.d.))

Longitudinal Data Structure

- Person-level dataset [wide-dataset]:
  - Each subject has one row [or record]
  - Repeated measures appear as additional variables
  - No explicit “time” variable
- Person-period dataset [long-dataset]:
  - A subject identifier
  - A time indicator
  - Outcome variable[s]
  - Predictor variable[s]

Example Data

- Data comes from the *National Youth Survey* (NYS; Raudenbush & Chan, 1992)
- Five waves, ages 11 - 15
- TOL, Tolerance of deviant behavior
  (1 = very wrong, 4 = not wrong at all)
- MALE, 1 for male, 0 for female
- EXP, self reported exposure to deviant behavior at age 11
  (0 =none, 4 =all).

“Person-level” Data Set

ID	TOL11	TOL12	TOL13	TOL14	TOL15	MALE	EXP
9	2.23	1.79	1.90	2.12	2.66	0	1.54
45	1.12	1.45	1.45	1.45	1.99	1	1.16
268	1.45	1.34	1.99	1.79	1.34	1	0.90
314	1.22	1.22	1.55	1.12	1.12	0	0.81
442	1.45	1.99	1.45	1.67	1.90	0	1.13
514	1.34	1.67	2.23	2.12	2.44	1	0.90
569	1.79	1.90	1.90	1.99	1.99	0	1.99
624	1.12	1.12	1.22	1.12	1.22	1	0.98
723	1.22	1.34	1.12	1.00	1.12	0	0.81
918	1.00	1.00	1.22	1.99	1.22	0	1.21
949	1.99	1.55	1.12	1.45	1.55	1	0.93
978	1.22	1.34	2.12	3.46	3.32	1	1.59
1105	1.34	1.90	1.99	1.90	2.12	1	1.38
1542	1.22	1.22	1.99	1.79	2.12	0	1.44
1552	1.00	1.12	2.23	1.55	1.55	0	1.04
1653	1.11	1.11	1.34	1.55	2.12	0	1.25

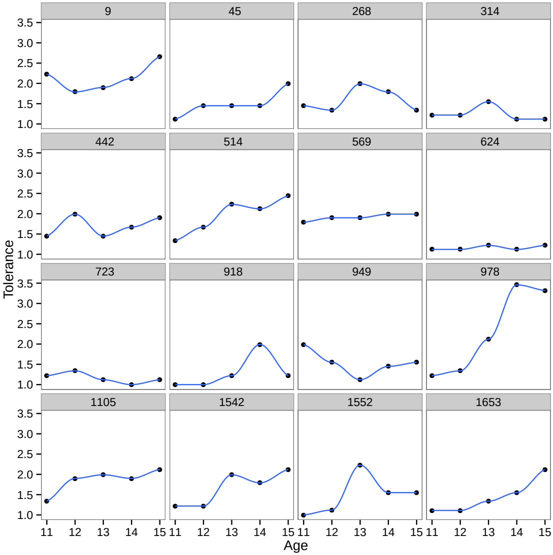
“Person-period” Data Set

ID	MALE	EXP	AGE	TOL
9	0	1.54	11	2.23
9	0	1.54	12	1.79
9	0	1.54	13	1.90
9	0	1.54	14	2.12
9	0	1.54	15	2.66
45	1	1.16	11	1.12
45	1	1.16	12	1.45
45	1	1.16	13	1.45
45	1	1.16	14	1.45
45	1	1.16	15	1.99
.	.	.	.	.
.	.	.	.	.
1653	0	1.25	11	1.11
1653	0	1.25	12	1.11
1653	0	1.25	13	1.34
1653	0	1.25	14	1.55
1653	0	1.25	15	2.12

Reshaping Data in R

```
library(foreign)

tol_dat <- read.dta("./data/nys.dta", convert.factors=T)
head(tol_dat)

library(reshape)

tol_long <- reshape(tol_dat,
  varying= c("tol11", "tol12", "tol13", "tol14", "tol15"),
  v.names= "tol",
  timevar= "age",
  times= c(11, 12, 13, 14, 15),
  direction= "long")

tol_long <- tol_long[order(tol_long$id),]
head(tol_long)
```

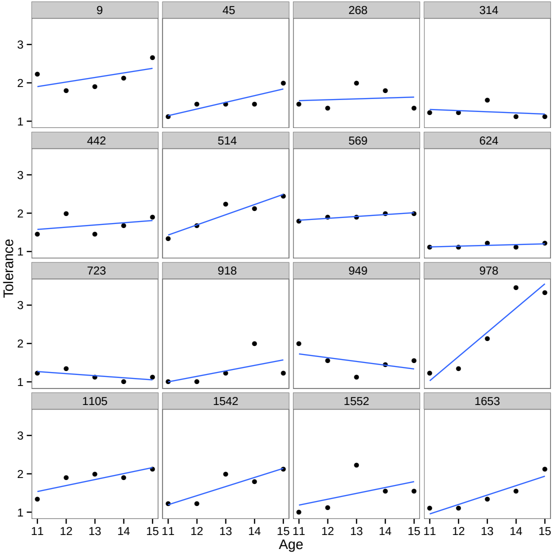
Exploring Longitudinal data: Non-parametric Summaries



Non-parametric Summary Plots in R

```
library(ggplot2)

p2 <- ggplot(data = tol_long, aes(y = tol, x = age))

p2 + geom_point() +
     stat_smooth(method=loess, se=F) +
     facet_wrap(~ id, ncol = 4) +
     theme_bw() +
     theme(panel.grid.major = element_blank(),
           panel.grid.minor = element_blank()) +
     ylab("Tolerance") +
     xlab("Age")
```

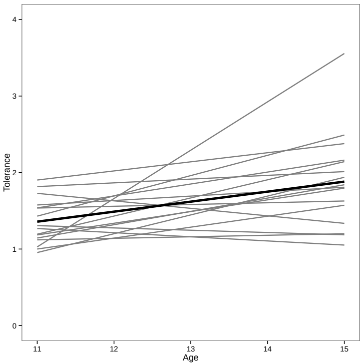
Exploring Longitudinal data: Fitted OLS Trajectories



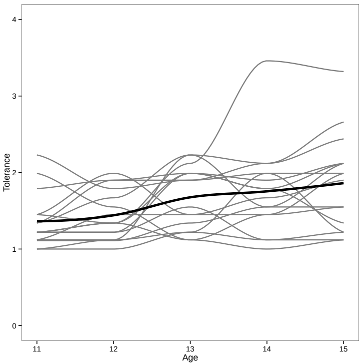
Fitted OLS Trajectory Plots in R

```
p2 <- ggplot(data = tol_long, aes(y = tol, x = age))

p2 + geom_point() +
     stat_smooth(method=lm, se=F) +
     facet_wrap(~ id, ncol = 4) +
     theme_bw() +
     theme(panel.grid.major = element_blank(),
           panel.grid.minor = element_blank()) +
     ylab("Tolerance") +
     xlab("Age")
```

Exploring Longitudinal data



(a) Fitted OLS

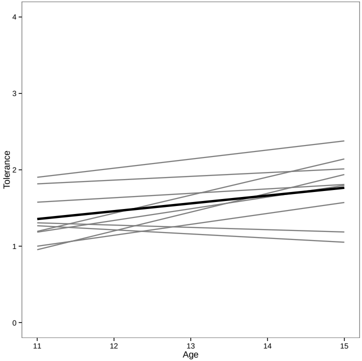


(b) Smooth non-parametric

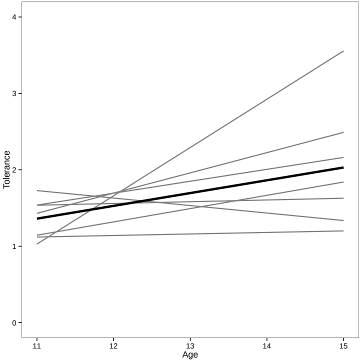
Exploratory Plots in R

```
p3 <- ggplot(data = tol_long, aes(y = tol, x = age))

p3 + stat_smooth(method=lm, aes(group= id), se=F, size=.75,
                 color="grey50") +
      stat_smooth(method=lm, se=F, size=1.5, color="black") +
      theme_bw() +
      theme(panel.grid.major = element_blank(),
            panel.grid.minor = element_blank()) +
      ylim(0, 4) + ylab("Tolerance") + xlab("Age")

p3 + stat_smooth(method=loess, aes(group= id), se=F, size=.75,
                 color="gray50") +
      stat_smooth(method=loess, se=F, size=1.5, color="black") +
      theme_bw() +
      theme(panel.grid.major = element_blank(),
            panel.grid.minor = element_blank()) +
      ylim(0, 4) + ylab("Tolerance") + xlab("Age")
```

Exploring Longitudinal data, by Male/Female



(c) Female

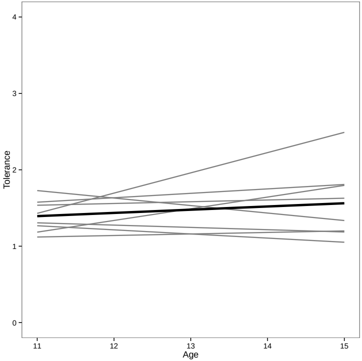


(d) Male

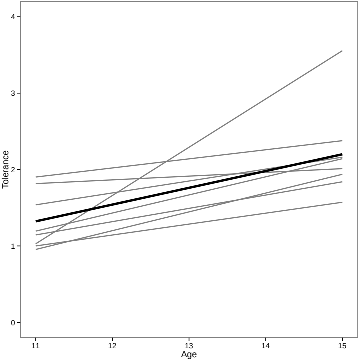
More Exploratory Plots in R!

```
p4_male <- ggplot(data = subset(tol_long, male==1),
                  aes(y = tol, x = age))

p4_male + stat_smooth(method=lm, aes(group= id), se=F, size=.75,
                      color="grey50") +
          stat_smooth(method=lm, se=F, size=1.5, color="black") +
          theme_bw() +
          theme(panel.grid.major = element_blank(),
                panel.grid.minor = element_blank()) +
          ylim(0, 4) + ylab("Tolerance") + xlab("Age")

p4_female <- ggplot(data = subset(tol_long, male==0),
                    aes(y = tol, x = age))

p4_female + stat_smooth(method=loess, aes(group= id), se=F, size=.75,
                       color="gray50") +
           stat_smooth(method=lm, se=F, size=1.5, color="black") +
           theme_bw() +
           theme(panel.grid.major = element_blank(),
                 panel.grid.minor = element_blank()) +
           ylim(0, 4) + ylab("Tolerance") + xlab("Age")
```

Exploring Longitudinal data, by Exposure (High > 1.145)



(e) Low Exposure



(f) High Exposure

Even More Exploratory Plots in R!!

```
tol_long$hiexp <- tol_long$exposure > 1.145

p4_hiexp <- ggplot(data = subset(tol_long, hiexp==T),
                   aes(y = tol, x = age))

p4_hiexp + stat_smooth(method=lm, aes(group= id), se=F, size=.75,
                       color="grey50") +
           stat_smooth(method=lm, se=F, size=1.5, color="black") +
           theme_bw() +
           theme(panel.grid.major = element_blank(),
                 panel.grid.minor = element_blank()) +
           ylim(0, 4) + ylab("Tolerance") + xlab("Age")

p4_loxp <- ggplot(data = subset(tol_long, hiexp!=T),
                  aes(y = tol, x = age))

p4_loexp + stat_smooth(method=lm, aes(group= id), se=F, size=.75,
                       color="gray50") +
          stat_smooth(method=loess, se=F, size=1.5, color="black") +
           theme_bw() +
           theme(panel.grid.major = element_blank(),
                 panel.grid.minor = element_blank()) +
           ylim(0, 4) + ylab("Tolerance") + xlab("Age")
```

The Multilevel Model for Change

The first example is limited to:

- Linear change model
- Time-structured data set
- Evaluation of one time-invariant dichotomous predictor

Example Data

- Data comes from Burchinal et al. (1997)
- 103 African-American infants born into low-income families
- At 6 months old, approximately half the sample ($n = 53$) were randomly assigned to participate in an intensive early intervention program designed to enhance cognitive functioning
- The remaining children ($n = 45$) were assigned to a control group
- Infants assessed 12 times between ages 6 and 96 months

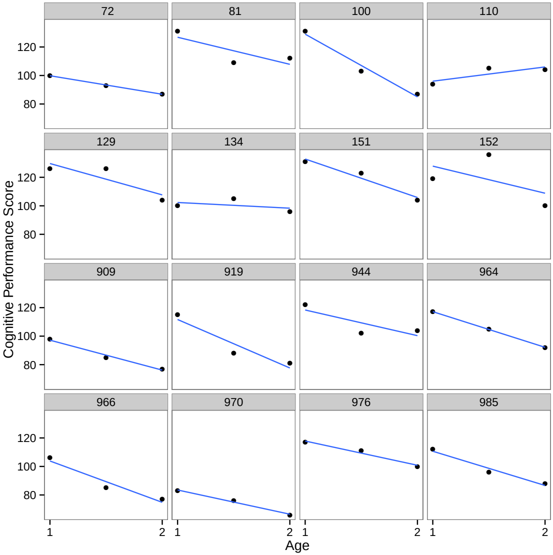
Example Data

- 3 waves of data–each child has three records
- *AGE* (in years) is the child's age at each assessment (1, 1.5, or 2)
- *COG* is the child's cognitive performance score at each assessment
- *PROGRAM* is a dichotomous covariate, 1= treatment and 0= control

Example Data

ID	COG	AGE	PROGRAM
68	103	1.0	1
68	119	1.5	1
68	96	2.0	1
70	106	1.0	1
70	107	1.5	1
70	96	2.0	1
.	.	.	.
.	.	.	.
984	106	1.0	0
984	89	1.5	0
984	99	2.0	0
985	112	1.0	0
985	96	1.5	0
985	88	2.0	0

Empirical Growth Plots: Fitted OLS Trajectories



The Multilevel Model for Change

$$
\begin{aligned}
Y_{ij} &= \pi_{0i} + \pi_{1i}(AGE_{ij} - 1) + \epsilon_{ij} \\
\pi_{0i} &= \gamma_{00} + \gamma_{01}(PROGRAM_i) + \zeta_{0i} \\
\pi_{1i} &= \gamma_{10} + \gamma_{11}(PROGRAM_i) + \zeta_{1i},
\end{aligned}
$$

$$
\epsilon_{ij} \sim N(0, \sigma_\epsilon^2) \text{ and } \begin{bmatrix} \zeta_{0i} \\ \zeta_{1i} \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_0^2 & \sigma_{01}^2 \\ \sigma_{10}^2 & \sigma_1^2 \end{bmatrix} \right)
$$

- The Level-1 Submodel
  - Describes how each person changes over time
  - Research questions about within-person change
- The Level-2 Submodel
  - Describes how these changes differ across people.
  - Research questions about how these changes vary across individuals

The Level-1 Submodel

$$Y_{ij} = [\pi_{0i} + \pi_{1i}(AGE_{ij} - 1)] + [\epsilon_{ij}]$$

Where,

- Y_{ij} represents the value of COG for child i at time j
  - i runs from 1 to 103
  - j runs from 1 to 3
- Brackets distinguish between the structural part and the stochastic part of the model
  - The structural part parallels the concept of "true score"
  - The stochastic part parallels the concept of "measurement error"

The Structural Part of the Level-1 Submodel

$$Y_{ij} = [\pi_{0i} + \pi_{1i}(AGE_{ij} - 1)] + [\epsilon_{ij}]$$

Our hypothesis about the shape of each subject's *true trajectory of change* over time

- π_{0i} represents child i's true cognitive performance at $X = 0$.
  - π_{01} is the intercept for child 1
  - π_{02} is the intercept for child 2
- π_{1i} represents the slope of the postulated individual change trajectory
  - If π_{1i} is positive, subject i's outcome increases over time

The Stochastic Part of the Level-1 Submodel

$$Y_{ij} = [\pi_{0i} + \pi_{1i}(AGE_{ij} - 1)] + [\epsilon_{ij}]$$

- ϵ_{ij} represents the effect of random error associated with individual i at time j
- ϵ_{ij} is unobserved so we must make assumptions about the distribution of level 1 residuals from occasion to occasion and from person to person.

The Stochastic Part of the Level-1 Submodel

$$\epsilon_{ij} \overset{iid}{\sim} N(0, \sigma_\epsilon^2)$$

- "Classical" assumptions specify residuals as independently and identically distributed ("iid"), with homoscedastic variance across occasions and individuals.
- Classical assumptions may not hold with longitudinal data as residuals may be autocorrelated and heteroscedastic over time (hold tight, more on this later).

The Level-2 Submodel

$$\pi_{0i} = [\gamma_{00} + \gamma_{01}(PROGRAM_i)] + [\zeta_{0i}]$$
$$\pi_{1i} = [\gamma_{10} + \gamma_{11}(PROGRAM_i)] + [\zeta_{1i}]$$

Where,

- π_{0i} and π_{1i} represents the level-1 change parameters–initial status and linear growth
- brackets distinguish between the structural part and the stochastic part of the model
  - the structural part parallels the concept of "true score"
  - the stochastic part parallels the concept of "measurement error"

The Structural Part of the Level-2 Submodel

$$\pi_{0i} = [\gamma_{00} + \gamma_{01}(PROGRAM_i)] + [\zeta_{0i}]$$

$$\pi_{1i} = [\gamma_{10} + \gamma_{11}(PROGRAM_i)] + [\zeta_{1i}]$$

Where,

- γs represent the level-2 regression parameters–known as *fixed effects*
- fixed effects capture inter individual differences in the true change trajectory
- interpret fixed effects as a *prototypical individual*:
  - γ_{00} represents the *average initial status* for children not enrolled in the treatment ($PROGRAM = 0$)
  - γ_{10} represents the *average annual growth* for children not enrolled in the treatment ($PROGRAM = 0$)
  - $\gamma_{00} + \gamma_{01}$ represents the *average initial status* for children enrolled in the treatment ($PROGRAM = 1$)
  - $\gamma_{10} + \gamma_{01}$ represents the *average annual growth* for children enrolled in the treatment ($PROGRAM = 1$)

The Stochastic Part of the Level-2 Submodel

$$\pi_{0i} = [\gamma_{00} + \gamma_{01}(PROGRAM_i)] + [\zeta_{0i}]$$
$$\pi_{1i} = [\gamma_{10} + \gamma_{11}(PROGRAM_i)] + [\zeta_{1i}]$$

Where,

- ζ represent the residuals–what remained *unexplained by the fixed effects*
- less interested in values of ζ than in the population summaries of the variances σ_0^2 and σ_1^2, and covariance σ_{01}^2

The Stochastic Part of the Level-2 Submodel

$$\begin{bmatrix} \zeta_{0i} \\ \zeta_{1i} \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_0^2 & \sigma_{01}^2 \\ \sigma_{10}^2 & \sigma_1^2 \end{bmatrix} \right)$$

Standard assumption about the level-2 residuals:

- Bivariate normal distribution
- With a mean of zero, and
- Unknown variance and covariance

R Function For Displaying Model Results

```
hlm.output <- function(x,npar=TRUE,print=F) {

cc <- fixef(x)
se <- sqrt(diag(vcov(x)))
coef.table <- round(cbind("Estimate"= cc,
 "Std.Err"= se,
 "z"= cc/se,
 "95CIL"= cc - (se * 1.96),
 "95CIU"= cc + (se * 1.96)),3)

var.cov <- as.data.frame(VarCorr(x))[,c(1, 2, 3, 4)]

mod_dev <- -2*(as.numeric(logLik(x)))
mod_df <- as.numeric(attr(logLik(x), "df"))
mod_bic <- AIC((ll <- logLik(x)), k = log(attr(ll,"nobs")))

mod_fit <- as.data.frame(c(mod_dev, mod_df, mod_bic),
row.names= c("Deviance", "df", "BIC"))
colnames(mod_fit) <- "Model Fit"

cat("\nFixed Effects\n"); print(coef.table)
cat("\nVarinace Components\n"); print(var.cov)
print(mod_fit)
}
```

Estimating a Multilevel Model for Change in R

```
library(lme4)

m3 <- lmer(cog ~ time + time*program + (time| id),
      data= ei_l,
      REML= F)

summary(m3)

hlm.output(m3)
```

Model Results

	Parameter	Estimate	ase	95% CI
Fixed Effects				
π_{0i}, Initial status	γ_{00}, Intercept	107.84	2.04	[103.85,111.83]
	γ_{01}, PROGRAM	6.86	1.88	[1.54,12.17]
π_{0i}, Rate of change	γ_{10}, Intercept	-21.13	1.88	[-24.83,-17.44]
	γ_{11}, PROGRAM	5.27	2.51	[0.35,10.19]
Variance Components				
Level 1:	σ_{ϵ}^2	74.76		
Level 2:	σ_0^2	123.97		
	σ_1^2	10.10		
	σ_{01}^2	-35.38		

Interpreting Fixed Effects

$$\hat{\pi}_{0i} = 107.84 + 6.86(PROGRAM_i)$$

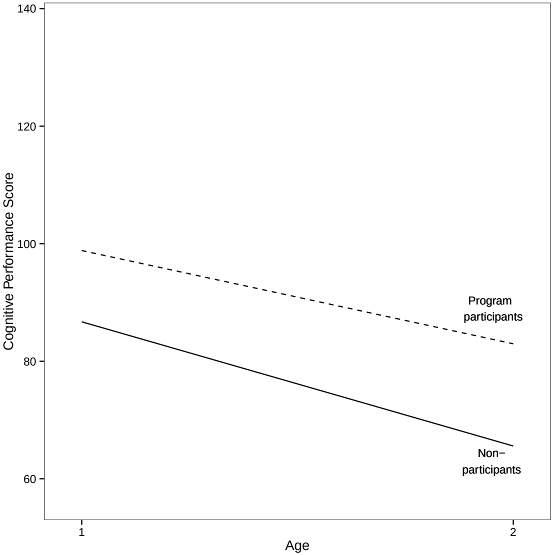
$$\hat{\pi}_{1i} = -21.13 + 5.27(PROGRAM_i)$$

Where,

- $107.84 =$ Initial status (*COG* at age=1) for the average *nonparticipant*
- $6.86 =$ Difference in initial status for the average *participant*
- $-21.13 =$ Annual rate of change for the average *nonparticipant*
- $5.27 =$ Difference in annual rate of change for the average *participant*

1. What is the estimated initial status for participants?
2. What is the estimated annual rate of change for participants?

Fitted Change Trajectories in COG



Plotting Change Trajectories In R

```
m3_fe <- fixef(m3)

prog0= function(x){
        m3_fe[1] + m3_fe[3]*(0) + m3_fe[2]*(x) + m3_fe[4]*(0)*(x)
        }
prog1= function(x){
        m3_fe[1] + m3_fe[3]*(1) + m3_fe[2]*(x) + m3_fe[4]*(1)*(x)
        }

tmp <- data.frame(x= min(ei_l$age) : max(ei_l$age))
```

Plotting Change Trajectories In R

```
m3_fit <- qplot(x, data=tmp)

m3_fit +
 stat_function(fun=prog1, linetype="dashed") +
 geom_text(aes(label="Program\nparticipants",
              x=1.95, y= 89, size=1)) +
 stat_function(fun=prog0) +
 geom_text(aes(label="Non-\nparticipants",
              x=1.95, y= 63, size=1)) +
 theme_bw() +
 theme(panel.grid.major = element_blank(),
       panel.grid.minor = element_blank()) +
 scale_x_continuous(breaks = c(1, 2), "Age") +
 scale_y_continuous(limits = c(min(ei_l$cog), max(ei_l$cog)),
                   "Cognitive Performance Score") +
 theme(legend.position="none")
```

Single Parameter Tests for Fixed Effects

Testing the statistical significance of fixed effects is similar to multiple regression where $H_0 : \gamma = 0$ and $H_1 : \gamma \neq 0$

Test this hypothesis for each fixed effect by computing a z-statistic:

$$z = \frac{\hat{\gamma}}{ase(\hat{\gamma})}$$

Interpreting Variance Components

$$\sigma_\epsilon^2 = 74.76$$

$$\begin{bmatrix} \sigma_0^2 & \sigma_{01}^2 \\ \sigma_{10}^2 & \sigma_1^2 \end{bmatrix} = \begin{bmatrix} 123.97 & -35.38 \\ -35.38 & 10.10 \end{bmatrix}$$

Where,

- Level-1 residual variance, σ_ϵ^2, summarizes within-person variability
- Level-2 variance components summarize between-person variability in change trajectories
- Single-parameter tests of significance for variance components can be highly inconsistent

Extending the Multilevel Model for Change

- The composite formulation
- Unconditional means model and unconditional growth model
- Model building strategies

Adolescent Alcohol Use Data

- Curran, Stice, and Chassin (1997) collected 3 waves of data
- Time-structured data set of 82 adolescents beginning at age 14.
  - *ALCUSE*, the level of alcohol consumption during the *previous* year
  - *AGE*, the age of the child at the time of data collection
  - *PEER*, a measure of alcohol use among the adolescent's peers
  - *COA*, a dichotomous covariate, indicating if the adolescent is a child of an alcoholic (1=yes, 0=no)

ALCUSE and *PEER* are generated by computing the square root of the sum of the participants' responses across each variable's constituent items.

Composite Specification of the Multilevel Model for Change

$$\begin{aligned} Y_{ij} =& \pi_{0i} + \pi_{1i} TIME_{ij} + \epsilon_{ij} \\ \pi_{0i} &= \gamma_{00} + \gamma_{01} COA_i + \zeta_{0i} \\ \pi_{1i} &= \gamma_{10} + \gamma_{11} COA_i + \zeta_{1i} \end{aligned}$$

$$\begin{aligned} Y_{ij} =& \pi_{0i} + \pi_{1i} TIME_{ij} + \epsilon_{ij} \\ =& (\gamma_{00} + \gamma_{01} COA_i + \zeta_{0i}) + (\gamma_{10} + \gamma_{11} COA_i + \zeta_{1i}) TIME_{ij} + \epsilon_{ij} \\ =& \gamma_{00} + \gamma_{10} TIME_{ij} + \gamma_{01} COA_i + \gamma_{11} (COA_i x TIME_{ij}) + \\ & \zeta_{0i} + \zeta_{1i} TIME_{ij} + \epsilon_{ij} \end{aligned}$$

The Unconditional Means Model

$$Y_{ij} = \gamma_{00} + \zeta_{0i} + \epsilon_{ij}$$

$$\epsilon_{ij} \sim N(0, \sigma_\epsilon^2) \text{ and } \zeta_{0i} \sim N(0, \sigma_0^2)$$

- Describes and partitions the outcome *variation*.
- Assumes the true individual change trajectory for person i is flat, sitting at elevation $\gamma_{00} + \zeta_{0i}$, or π_{0i}.
- Average (*grand mean*) elevation, across everyone, is γ_{00}.
- Partions the total outcome variation by within-person, σ_ϵ^2 and between-person, σ_0^2.

Estimating the Unconditional Means Model in R

```
alc_m1 <- lmer(alcuse ~ 1 + (1|id),
          data= alc_l)
          REML= F)

summary(alc_m1)

hlm.output(alc_m1)
```

Model Results

	Model 1	Model 2	Model 3	Model 4	Model 5
Fixed Effects					
γ_{00}, Initial status	0.922	0.651	0.316	-0.317	-0.314
	(0.096)	(0.105)	(0.131)	(0.148)	(0.146)
γ_{01}, COA		1.88	0.743	0.579	0.571
			(0.195)	(0.162)	(0.146)
γ_{02}, PEER				0.694	0.695
				(0.112)	(0.111)
γ_{10}, Rate of change		0.271	0.293	0.429	0.425
		(0.062)	(0.084)	(0.114)	(0.106)
γ_{11}, COA			-0.049	0.014	
			0.125	(0.125)	
γ_{12}, PEER				-0.150	-0.151
				(0.086)	(0.085)
Variance Components					
σ^2_ϵ	0.562	0.337	0.337	0.337	0.337
σ^2_0	0.564	0.624	0.488	0.241	0.241
σ^2_1		0.151	0.151	0.139	0.139
σ^2_{01}		-0.068	-0.059	-0.006	-0.006

The Intraclass Correlation Coefficient, ICC

$$\rho = \frac{\sigma_0^2}{\sigma_0^2 + \sigma_\epsilon^2}$$

$$= \frac{0.562}{0.562 + 0.564} = \frac{0.562}{1.126} = 0.501$$

```
lvl2.var <- as.data.frame(VarCorr(alc_m1))[1,4]
lvl1.var <- as.data.frame(VarCorr(alc_m1))[2,4]
icc <- lvl2.var / (lvl2.var + lvl1.var)
```

- Describes the proportion of total variance that lies between people.
- Also know as the *error autocorrelation coefficient.*

The Unconditional Growth Model

$$Y_{ij} = \gamma_{00} + \gamma_{10} TIME_{ij} + \zeta_{0i} + \zeta_{1i} TIME_{ij} + \epsilon_{ij}$$

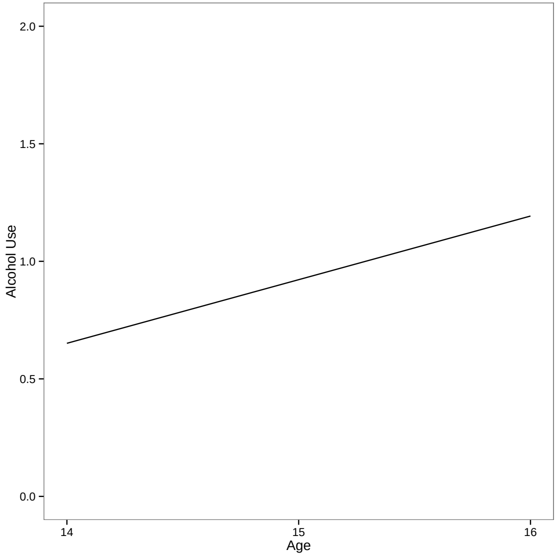
$$\epsilon_{ij} \sim N(0, \sigma^2_\epsilon) \text{ and } \begin{bmatrix} \zeta_{0i} \\ \zeta_{1i} \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma^2_0 & \sigma_{01} \\ \sigma_{10} & \sigma^2_1 \end{bmatrix} \right)$$

- Describes the unconditional initial status and rate of change for the population.
- $\gamma_{00} + \zeta_{0i}$ represents the interindividual initial status
- $\gamma_{10} + \zeta_{1i}$ represents the interindividual rate of change
- σ^2_ϵ summarizes each person's data around his/her linear change trajectory
- σ^2_0 and σ^2_0 summarize between-person variability in initial status and rates of change.

Estimating the Unconditional Growth Model

```
alc_m2 <- lmer(alcuse ~ age_14 + (age_14|id),
          data=alc_l,
          REML=F)

summary(alc_m2)

hlm.output(alc_m2)
```

The Unconditional Growth Model Graphically



Pseudo R^2 – Understanding the effect of *TIME*

$$\frac{\sigma^2_{\epsilon_{Model1}} - \sigma^2_{\epsilon_{Model2}}}{\sigma^2_{\epsilon_{Model1}}} = \frac{0.562 - 0.337}{0.562} = 0.4004$$

- 40% of the with-in person variation in *ALCUSE* is systematically associated with linear *TIME*.

The Unconditional Growth Model Covariance

$$\begin{bmatrix} \zeta_{0i} \\ \zeta_{1i} \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_0^2 & \sigma_{01} \\ \sigma_{10} & \sigma_1^2 \end{bmatrix} \right)$$

$$\hat{\rho}_{\pi_0\pi_1} = \hat{\rho}_{01} = \frac{\sigma_{01}}{\sqrt{\sigma_0^2\sigma_0^2}} = \frac{-0.068}{\sqrt{(0.624)(0.151)}} = -0.22$$

- The linear relationship between *ALCUSE* at age 14, γ_{00} and rate of change in *ALCUSE* between age 14 and 16, γ_{10} is weakly negative.

A Taxonomy Of Statistical Models

- A *taxonomy* of models is a "systematic sequence of models that, as a set, address your research question" (Singer & Willett, 2003, p. 105).
- Distinguish between *control* predictors and *question* predictors.
  - In our example, we will assume our research questions focuses on *COA*.
  - *PEER* is used as a control.

The Uncontrolled Effects of *COA*

$$
\begin{aligned}
Y_{ij} =& \pi_{0i} + \pi_{1i} TIME_{ij} + \epsilon_{ij} \\
\pi_{0i} &= \gamma_{00} + \gamma_{01} COA_i + \zeta_{0i} \\
\pi_{1i} &= \gamma_{10} + \gamma_{10} COA_i + \zeta_{1i}
\end{aligned}
$$

$$
Y_{ij} = \gamma_{00} + \gamma_{01} COA_i + \gamma_{10} TIME_{ij} + \gamma_{10}(TIME_{ij} COA_i) + \zeta_{0i} + \zeta_{1i} TIME_{ij} + \epsilon_{ij}
$$

$$
\epsilon_{ij} \sim N(0, \sigma_\epsilon^2) \text{ and } \begin{bmatrix} \zeta_{0i} \\ \zeta_{1i} \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_0^2 & \sigma_{01} \\ \sigma_{10} & \sigma_1^2 \end{bmatrix} \right)
$$

- γ_{01} describes the difference in the level of *ALCUSE* at age 14 for children with and without alcoholic parents.
- γ_{11} describes the impact of *COA* on the rate of change in *ALCUSE* between ages 14 and 16.

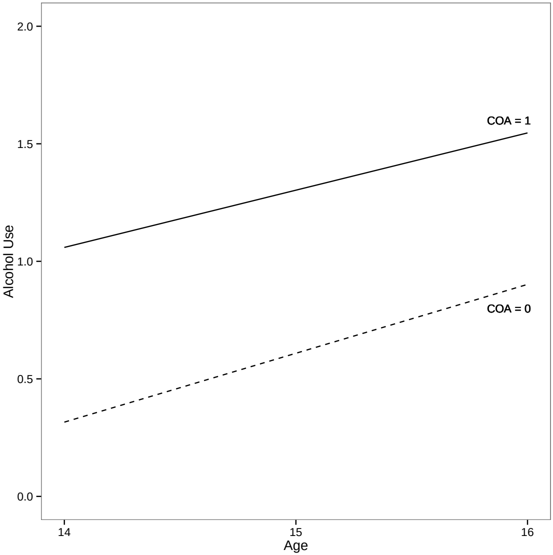
Estimating the Uncontrolled Effects of *COA* in R

```
alc_m3 <- lmer(alcuse ~ age_14*coa + (age_14|id),
           data=alc_l,
           REML=F)

summary(alc_m3)

hlm.output(alc_m3)
```

The Uncontrolled Effects of *COA*

	Model 1	Model 2	Model 3	Model 4	Model 5
Fixed Effects					
γ_{00}, Initial status	0.922	0.651	0.316	−0.317	−0.314
	(0.096)	(0.105)	(0.131)	(0.148)	(0.146)
γ_{01}, COA		1.88	0.743	0.579	0.571
			(0.195)	(0.162)	(0.146)
γ_{02}, PEER				0.694	0.695
				(0.112)	(0.111)
γ_{10}, Rate of change		0.271	0.293	0.429	0.425
		(0.062)	(0.084)	(0.114)	(0.106)
γ_{11}, COA			−0.049	0.014	
			0.125	(0.125)	
γ_{12}, PEER				−0.150	−0.151
				(0.086)	(0.085)
Variance Components					
σ_ϵ^2	0.562	0.337	0.337	0.337	0.337
σ_0^2	0.564	0.624	0.488	0.241	0.241
σ_1^2		0.151	0.151	0.139	0.139
σ_{01}		−0.068	−0.059	−0.006	−0.006

The Uncontrolled Effects of *COA* Graphically



The Uncontrolled Effects of *COA*

	Model 1	Model 2	Model 3	Model 4	Model 5
Pseudo R^2 Statistics					
R^2_ϵ		0.400	0.000	0.000	0.000
R^2_0			0.219	0.501	0.000
R^2_1			0.004	0.076	0.000
Goodness-of-fit					
Deviance	670.156	636.611	621.203	588.691	588.703
df	3	6	8	10	9
BIC	686.672	669.643	665.245	643.744	638.251

Pseudo R^2 in R

```
rsq.output <- function(x, y, npar=TRUE,print=F) {
options(scipen=999)

...

m1_int_var<- as.data.frame(VarCorr(x))
m1_var_0 <- m1_int_var$vcov[which
          (m1_int_var[,"var1"]=="(Intercept)" &
          is.na(m1_int_var[,"var2"]))]

m2_int_var<- as.data.frame(VarCorr(y))
m2_var_0 <- m2_int_var$vcov[which
          (m2_int_var[,"var1"]=="(Intercept)" &
          is.na(m2_int_var[,"var2"]))]

m2_r2_0 <- (m1_var_0 - m2_var_0) / m1_var_0 ## R2_0

...

}
```

Comparing Models Using Deviance Statistics

- Comparing models using *deviance statistics* is a more robust approach than using single parameter tests
    1. Superior statistical properties.
    2. Permits composite tests on several parameters.
    3. "Reserves the reservoir of Type I error" (Singer & Willett, 2003, p. 116).
- FML tests all parameters while RML tests only variance components.

$$\text{Deviance} = -2[LL_{\text{current model}} - LL_{\text{saturated model}}]$$

- *LL* is the log-likelihood, a byproduct of ML estimation–the larger the LL (closer to 0) the better the fit.
- The saturated model is a general mode that fits the data perfectly.
- Deviance quantifies how much worse the current model fits the data compared to the best possible model.

Comparing Models Using Deviance Statistics

$$\begin{aligned}
\text{Deviance} &= -2[LL_{\text{current model}} - LL_{\text{saturated model}}] \\
&= -2[LL_{\text{current model}} - 0] \\
&= -2LL_{\text{current model}}
\end{aligned}$$

- $LL_{\text{saturated model}} = 0$ because the probability that the model will perfectly fit the data is 1 ($log(1) = 0$).
- -2 because standard normal theory assumptions say that comparing nested models has a known distribution.

Comparing Models Using Deviance Statistics

Deviance-based Hypothesis Tests:

- Data set must be unchanged across models.
- The former model must be nested within the latter model.
- Compute the number of additional constraints imposed.
- ΔD is distributed asymptotically as a χ^2 distribution. with $d.f. =$ the number of independent constraints imposed.

$$\Delta D = \text{Deviance}_{\text{Reduced Model}} - \text{Deviance}_{\text{Full Model}}$$

$$\begin{aligned}\Delta D &= \text{Deviance}_{\text{Model 2}} - \text{Deviance}_{\text{Model 3}} \\ &= 636.611 - 621.203 = 15.408\end{aligned}$$

15.408 exceeds the χ^2 .001 critical value at 2 $d.f.$ (13.816), allowing us to reject the null hypothesis that γ_{01} and γ_{11} are simultaneously 0.

Deviance Test Function in R

```
dev <- function(a, b){
return(1 - pchisq(
         m1_dev <- -2*(as.numeric(logLik(a))) -
                   -2*(as.numeric(logLik(b))),
     as.numeric(attr(logLik(b), "df"))-
     as.numeric(attr(logLik(a), "df"))))
}
```

The Controlled Effects of *COA*

$$
\begin{aligned}
Y_{ij} = &\gamma_{00} + \gamma_{01} COA_i + \gamma_{02} PEER_i + \gamma_{10} TIME_{ij} + \\
&\gamma_{11}(TIME_{ij} COA_i) + \gamma_{12}(TIME_{ij} PEER_i) + \zeta_{0i} + \zeta_{1i} TIME_{ij} + \epsilon_{ij}
\end{aligned}
$$

$$
\epsilon_{ij} \sim N(0, \sigma_\epsilon^2) \text{ and } \begin{bmatrix} \zeta_{0i} \\ \zeta_{1i} \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_0^2 & \sigma_{01} \\ \sigma_{10} & \sigma_1^2 \end{bmatrix} \right)
$$

- γ_{02} describes the impact of peer alcohol use on the level of *ALCUSE* at age 14 for children, controlling for *COA*.
- γ_{12} describes the impact of peer alcohol use on the rate of change in *ALCUSE* between ages 14 and 16, controlling for *COA*..

Estimating the Controlled Effects of *COA*

```
alc_m4 <- lmer(alcuse ~ age_14*coa + age_14*peer + (age_14|id),
           data=alc_l,
           REML=F)

summary(alc_m4)

hlm.output(alc_m4)
```

The Controlled Effects of *COA*

	Model 1	Model 2	Model 3	Model 4	Model 5
Fixed Effects					
γ_{00}, Initial status	0.922	0.651	0.316	−0.317	−0.314
	(0.096)	(0.105)	(0.131)	(0.148)	(0.146)
γ_{01}, COA		1.88	0.743	0.579	0.571
			(0.195)	(0.162)	(0.146)
γ_{02}, PEER				0.694	0.695
				(0.112)	(0.111)
γ_{10}, Rate of change		0.271	0.293	0.429	0.425
		(0.062)	(0.084)	(0.114)	(0.106)
γ_{11}, COA			−0.049	0.014	
			0.125	(0.125)	
γ_{12}, PEER				−0.150	−0.151
				(0.086)	(0.085)
Variance Components					
σ^2_ϵ	0.562	0.337	0.337	0.337	0.337
σ^2_0	0.564	0.624	0.488	0.241	0.241
σ^2_1		0.151	0.151	0.139	0.139
σ_{01}		−0.068	−0.059	−0.006	−0.006

The Controlled Effects of *COA*

	Model 1	Model 2	Model 3	Model 4	Model 5
Pseudo R^2 Statistics					
R^2_ϵ		0.400	0.000	0.000	0.000
R^2_0			0.219	0.501	0.000
R^2_1			0.004	0.076	0.000
Goodness-of-fit					
Deviance	670.156	636.611	621.203	588.691	588.703
df	3	6	8	10	9
BIC	686.672	669.643	665.245	643.744	638.251

Final Model for Controlled Effects of *COA*

$$\begin{aligned} Y_{ij} = &\gamma_{00} + \gamma_{01} COA_i + \gamma_{02} PEER_i + \gamma_{10} TIME_{ij} + \\ &\gamma_{12}(TIME_{ij} PEER_i) + \zeta_{0i} + \zeta_{1i} TIME_{ij} + \epsilon_{ij} \end{aligned}$$

$$\epsilon_{ij} \sim N(0, \sigma_\epsilon^2) \text{ and } \begin{bmatrix} \zeta_{0i} \\ \zeta_{1i} \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_0^2 & \sigma_{01} \\ \sigma_{10} & \sigma_1^2 \end{bmatrix} \right)$$

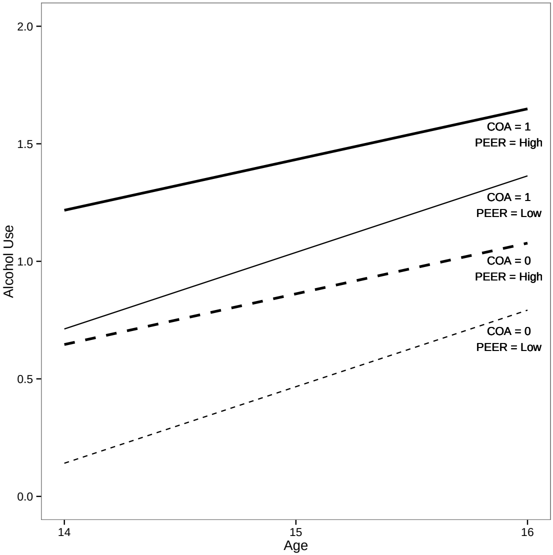
Estimating the Final Model

```
alc_m5 <- lmer(alcuse ~ age_14 + coa + age_14*peer + (age_14|id),
data=alc_l,
REML=F)

summary(alc_m5)

hlm.output(alc_m5)
```

The Controlled Effects of *COA*

	Model 1	Model 2	Model 3	Model 4	Model 5
Fixed Effects					
γ_{00}, Initial status	0.922	0.651	0.316	-0.317	-0.314
	(0.096)	(0.105)	(0.131)	(0.148)	(0.146)
γ_{01}, COA		1.88	0.743	0.579	0.571
			(0.195)	(0.162)	(0.146)
γ_{02}, PEER				0.694	0.695
				(0.112)	(0.111)
γ_{10}, Rate of change		0.271	0.293	0.429	0.425
		(0.062)	(0.084)	(0.114)	(0.106)
γ_{11}, COA			-0.049	0.014	
			0.125	(0.125)	
γ_{12}, PEER				-0.150	-0.151
				(0.086)	(0.085)
Variance Components					
σ_ϵ^2	0.562	0.337	0.337	0.337	0.337
σ_0^2	0.564	0.624	0.488	0.241	0.241
σ_1^2		0.151	0.151	0.139	0.139
σ_{01}		-0.068	-0.059	-0.006	-0.006

The Controlled Effects of *COA* Graphically



The Controlled Effects of *COA*

	Model 1	Model 2	Model 3	Model 4	Model 5
Pseudo R^2 Statistics					
R^2_ϵ		0.400	0.000	0.000	0.000
R^2_0			0.219	0.501	0.000
R^2_1			0.004	0.076	0.000
Goodness-of-fit					
Deviance	670.156	636.611	621.203	588.691	588.703
df	3	6	8	10	9
BIC	686.672	669.643	665.245	643.744	638.251

Deviance Tests When Model Trimming

$$\Delta D = \text{Deviance}_{\text{Reduced Model}} - \text{Deviance}_{\text{Full Model}}$$

$$\begin{aligned}\Delta D &= \text{Deviance}_{\text{Model 5}} - \text{Deviance}_{\text{Model 4}}\\ &= 588.703 - 588.691 = 0.012\end{aligned}$$

0.012 does not exceed the χ^2 .001 critical value at 1 $d.f.$ (3.841). We are unable to reject the null hypothesis that γ_{11} is 0.

AIC and BIC

- AIC: Akaike Information Criterion (Akaike, 1973)
  - scale factor $= 1$
  - number of parameters (fixed effects and variance components)
- BIC: Bayesian Information Criterion (Schwarz, 1978)
  - scale factor $= .5(\log(N))$
  - number of parameters (fixed effects and variance components)

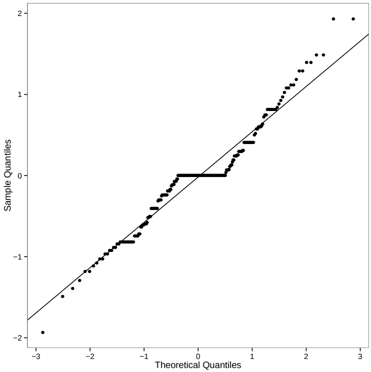
$$\begin{aligned}\text{IC} &= -2[LL - (\text{scale factor})(\text{number of parameters in the model})] \\ &= \text{Deviance} + 2(\text{scale factor})(\text{number of parameters in the model})\end{aligned}$$

```
mod_bic <- AIC((ll <- logLik(x)), k = log(attr(ll,"nobs")))
```

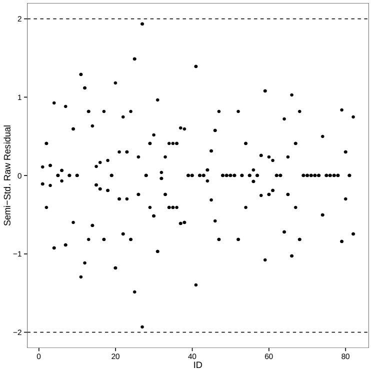
Evaluating a Model's Assumptions

- Checking functional form
  - At level 1: OLS-estimated individual change trajectories.
  - At level 2: OLS-estimates of the individual growth parameters.
- Checking normality
  - Normal probability plots
  - Standardized residuals vs. ID
- Homoscedasticity
  - Level 1 residuals against level 1 predictors
  - Level 2 residuals against level 2 predictors

Level-1 Normality Assumption

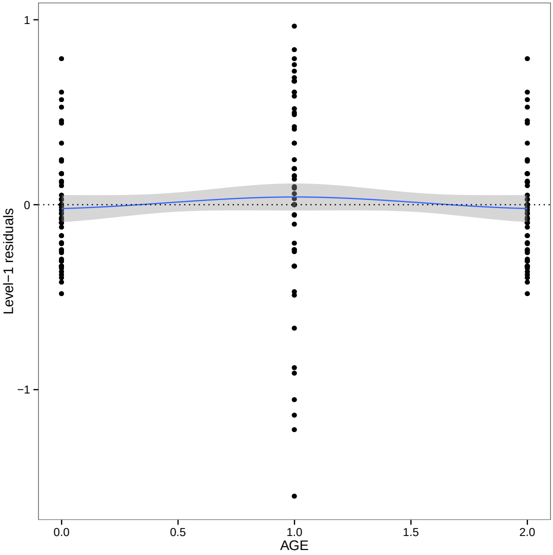


(g) Normal probability plot



(h) Semi-std. Residuals vs. ID

Level-1 Homoscedasticity Assumption



Extending the Multilevel Model for Change

- Variably spaced measurements.
- Varying number of measurements.
- Time-varying covariates.
- Re-centering the effect of time.
- Non-linear growth (just a taste)

Time-structured vs. Unstructured Data

ID	WAVE	AGEGRP	AGE	PIAT
04	1	6.5	6.00	18
04	2	8.5	8.50	31
04	3	10.5	10.67	50
27	1	6.5	6.25	19
27	2	8.5	9.17	36
27	3	10.5	10.92	57
31	1	6.5	6.33	18
31	2	8.5	8.83	31
31	3	10.5	10.92	51
.	.	.	.	.
.	.	.	.	.

Centering TIME variables in R

```
read_l <- read.dta("./data/reading_pp.dta", convert.factors=T)

centered <- read_l[ ,3:4] - min(read_l$agegrp)

dimnames(centered)[[2]] <- c("agegrp.c", "age.c")

read_l <- cbind(read_l, centered)
```

Estimating Structured and Unstructured TIME in R

```
# forcing structure on data
agegrp <- lmer(piat ~ agegrp.c + (agegrp.c | id),
        read_l,
        REML = F)

hlm.output(agegrp)

# using unstructured data
age <- lmer(piat ~ age.c + (age.c | id),
        read_l,
        REML = F)

hlm.output(age)
```

Time-structured vs. Unstructured Data: Results

	AGEGRP-6.5	AGE-6.5
Fixed Effects		
γ_{00}, Initial status	21.163	21.061
	(0.614)	(0.559)
γ_{10}, Rate of change	5.031	4.540
	(0.296)	(0.261)
Variance Components		
σ_ϵ^2	27.043	27.447
σ_0^2	11.046	5.107
σ_1^2	4.397	3.301
σ_{01}^2	1.647	2.3667
Goodness-of-fit		
Deviance	1819.949	1803.896
df	6	6
BIC	1853.473	1837.419

Time-structured vs. Unstructured Data

- γ_{10} is half a point larger for *AGEGRP* (5.031 vs. 4.540)
- σ_0^2 and σ_1^2 are much larger for (AGEGRP)
- BIC is smaller for *AGE* indicating a better fit

Lesson: Never force an unstructured data set to be structured.

Unbalanced Data

- All subjects can contribute to a multilevel model regardless of how many waves of data they contribute.
- As long as there are enough subjects with enough data points, the model should estimate, given the complexity of the model.
- Potential problems include:
  - Nonconvergence
  - Variance components may exceed boundary constraints (e.g., negative variance components)

Understanding the source of imbalance is critical (addressed in *Workshop 3: Longitudinal Data Analysis with Incomplete Data November 1, 9:00am - 3:00pm*)

Time-varying Covariates, Example Data

- Data comes from Ginexi et al. (2000)
- 254 participants who are in their first two months of job loss.
- Follow-up interviews conducted between 3 and 8 months and 10 and 16 months after job loss
- Center for Epidemiologic Studies' Depression (CES-D) scale.

Time-varying Covariates, Example Data

- 4 waves of data–each subject has four records
- *MONTHS* number of months since study began
- $CES - D$, $0 =$ low to $8 =$ serious distress
- *UNEMP* time-varying covariate $1 =$ unemployed, $0 =$ employed

Unconditional Growth Model

$$Y_{ij} = \gamma_{00} + \gamma_{10} MONTHS_{ij} + \zeta_{0i} + \zeta_{1i} MONTHS_{ij} + \epsilon_{ij}$$

$$\epsilon_{ij} \sim N(0, \sigma_{\epsilon}^2) \text{ and } \begin{bmatrix} \zeta_{0i} \\ \zeta_{1i} \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_0^2 & \sigma_{01} \\ \sigma_{10} & \sigma_1^2 \end{bmatrix} \right)$$

- γ_{00} is the average initial status CES-D score.
- γ_{10} is the average monthly rate of change in CES-D sores.

Model Results

	Model 1	Model 2	Model 3	Model 4
Fixed Effects				
γ_{00}, Initial status	17.669	12.666	11.125	11.198
	(0.776)	(1.242)	(0.901)	(0.793)
γ_{01}, MONTHS	−0.422	−0.202		
	(0.083)	(0.093)		
γ_{10}, UNEMP		5.111	7.000	6.924
		(0.989)	(0.920)	(0.933)
γ_{11}, UNEMP by MONTHS			−0.300	−0.303
			(0.107)	(0.112)
Variance Components				
σ_ϵ^2	68.850	62.388	69.857	59.097
σ_0^2	86.8489	93.518	67.294	45.765
σ_1^2	0.355	0.466		
σ_2^2				45.915
σ_3^2			0.296	0.764
Goodness-of-fit				
Deviance	5133.137	5107.603	5110.798	5095.584
df	6	7	7	10
BIC	5172.217	5153.196	5156.390	5160.672

Including a Time-varying Covariate *UNEMP*

$$Y_{ij} = \gamma_{00} + \gamma_{10} MONTHS_{ij} + \gamma_{20} UNEMP_{ij} + \zeta_{0i} + \zeta_{1i} MONTHS_{ij} + \epsilon_{ij}$$

$$\epsilon_{ij} \sim N(0, \sigma_\epsilon^2) \text{ and } \begin{bmatrix} \zeta_{0i} \\ \zeta_{1i} \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_0^2 & \sigma_{01} \\ \sigma_{10} & \sigma_1^2 \end{bmatrix} \right)$$

- γ_{10} is the average monthly rate of change in CES-D score, controlling for unemployment status.
- γ_{20} is the average difference, over time, in CES-D sores between the unemployed (0) and employed (1)

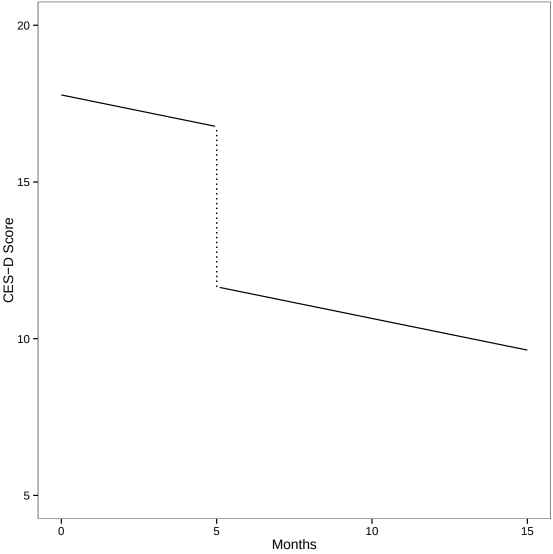
Estimating Time-varying Covariates in R

```
unemp_l <- read.dta("./data/unemploy.dta", convert.factors=T)

unemp_m2 <- lmer(cesd ~ months + unemp + (months | id),
            unemp_l,
            REML = FALSE)

summary(unemp_m2)
hlm.output(unemp_m2)
```

Model Results

	Model 1	Model 2	Model 3	Model 4
Fixed Effects				
γ_{00}, Initial status	17.669	12.666	11.125	11.198
	(0.776)	(1.242)	(0.901)	(0.793)
γ_{01}, MONTHS	−0.422	−0.202		
	(0.083)	(0.093)		
γ_{10}, UNEMP		5.111	7.000	6.924
		(0.989)	(0.920)	(0.933)
γ_{11}, UNEMP by MONTHS			−0.300	−0.303
			(0.107)	(0.112)
Variance Components				
σ_{ϵ}^2	68.850	62.388	69.857	59.097
σ_0^2	86.8489	93.518	67.294	45.765
σ_1^2	0.355	0.466		
σ_2^2				45.915
σ_3^2			0.296	0.764
Goodness-of-fit				
Deviance	5133.137	5107.603	5110.798	5095.584
df	6	7	7	10
BIC	5172.217	5153.196	5156.390	5160.672

Model-based Trajectory, Gaining Employment at Month 5



Plotting Model-based Trajectories in R

```
tmp1 <- data.frame(x= seq(0, 15, by=5))

unemp2_fe <- fixef(unemp_m2)

unemp2= function(x){
        ifelse(x <= 5 ,
        unemp2_fe[1] + unemp2_fe[3]*(1) + unemp2_fe[2]*(x) ,
        NA)
        }
unemp3= function(x){
        ifelse(x >= 5 ,
        unemp2_fe[1] + unemp2_fe[3]*(0) + unemp2_fe[2]*(x) ,
        NA)
        }
```

Plotting Model-based Trajectories in R

```
unemp1_fit <- ggplot(data=tmp1, aes(x))

unemp1_fit +
      stat_function(fun=unemp2) +
      stat_function(fun=unemp3) +
     geom_segment(aes(x= 5, y= 11.65568 , xend = 5, yend =  16.76699),
             linetype="dotted") +
      theme_bw() +
      theme(panel.grid.major = element_blank(),
            panel.grid.minor = element_blank()) +
      scale_x_continuous(limits=c(0, 15), "Months") +
      scale_y_continuous(limits=c(5, 20), "CES-D Score") +
      theme(legend.position="none")
```

Time-varying Predictor *UNEMP* by *TIME*

$$
\begin{aligned}
Y_{ij} =& \gamma_{00} + \gamma_{20} UNEMP_{ij} + \gamma_{30}(UNEMP_{ij} MONTHS_{ij}) + \\
& \zeta_{0i} + \zeta_{3i}(UNEMP_{ij} TIME_{ij}) + \epsilon_{ij}
\end{aligned}
$$

$$
\epsilon_{ij} \sim N(0, \sigma_\epsilon^2) \text{ and } \begin{bmatrix} \zeta_{0i} \\ \zeta_{3i} \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_0^2 & \sigma_{03} \\ \sigma_{30} & \sigma_3^2 \end{bmatrix} \right)
$$

- γ_{00} represents the flat trajectory in CES-D scores for employed.
- γ_{20} represents the difference in initial status in CES-D scores between the unemployed and employed.
- γ_{30} represents the average monthly rate if change in CES-D scores for the unemployed.

Estimating Time-varying Covariates in R

```
unemp_m4 <- lmer(cesd ~ unemp + unemp:months + (unemp:months | id),
                 unemp_l,
                 REML = FALSE)

summary(unemp_m4)

hlm.output(unemp_m4)
```

Model Results

	Model 1	Model 2	Model 3	Model 4
Fixed Effects				
γ_{00}, Initial status	17.669	12.666	11.125	11.198
	(0.776)	(1.242)	(0.901)	(0.793)
γ_{01}, MONTHS	−0.422	−0.202		
	(0.083)	(0.093)		
γ_{10}, UNEMP		5.111	7.000	6.924
		(0.989)	(0.920)	(0.933)
γ_{11}, UNEMP by MONTHS			−0.300	−0.303
			(0.107)	(0.112)
Variance Components				
σ_{ϵ}^2	68.850	62.388	69.857	59.097
σ_0^2	86.8489	93.518	67.294	45.765
σ_1^2	0.355	0.466		
σ_2^2				45.915
σ_3^2			0.296	0.764
Goodness-of-fit				
Deviance	5133.137	5107.603	5110.798	5095.584
df	6	7	7	10
BIC	5172.217	5153.196	5156.390	5160.672

Final Model

$$\begin{aligned} Y_{ij} = &\gamma_{00} + \gamma_{20} UNEMP_{ij} + \gamma_{20}(UNEMP_{ij} MONTHS_{ij}) + \\ &\zeta_{0i} + \zeta_{2i} UNEMP_{ij} + \zeta_{3i}(UNEMP_{ij} TIME_{ij}) + \epsilon_{ij} \end{aligned}$$

$$\epsilon_{ij} \sim N(0, \sigma_\epsilon^2) \text{ and } \begin{bmatrix} \zeta_{0i} \\ \zeta_{2i} \\ \zeta_{3i} \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_0^2 & \sigma_{02} & \sigma_{03} \\ \sigma_{20} & \sigma_2^2 & \sigma_{23} \\ \sigma_{30} & \sigma_{32} & \sigma_3^2 \end{bmatrix} \right)$$

- σ_0^2 is the variance in CES-D scores for employed subjects.
- σ_2^2 is the variance in CES-D scores for unemployed subjects at the first interview.
- σ_3^2 is the variance in average monthly rate of change in CES-D scores for unemployed subjects.

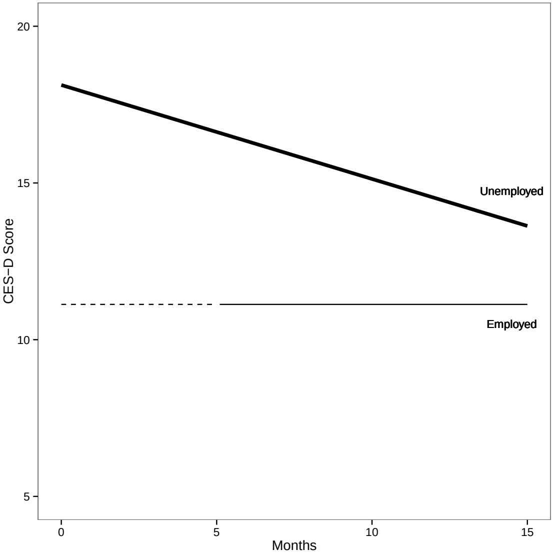
Estimating the Final Model in R (or this is how you would if...)

```
## This model is under identified.  See ALDA p. 163 for results.
unemp_m5 <- lmer(cesd ~ unemp + unemp:months +
                 (unemp + unemp:months | id),
                 unemp_l,
                 REML = FALSE,
                 control=lmerControl(optimizer="bobyqa"))

summary(unemp_m5)

hlm.output(unemp_m5)
```

Model Results

	Model 1	Model 2	Model 3	Model 4
Fixed Effects				
γ_{00}, Initial status	17.669	12.666	11.125	11.198
	(0.776)	(1.242)	(0.901)	(0.793)
γ_{01}, MONTHS	−0.422	−0.202		
	(0.083)	(0.093)		
γ_{10}, UNEMP		5.111	7.000	6.924
		(0.989)	(0.920)	(0.933)
γ_{11}, UNEMP by MONTHS			−0.300	−0.303
			(0.107)	(0.112)
Variance Components				
σ^2_ϵ	68.850	62.388	69.857	59.097
σ^2_0	86.8489	93.518	67.294	45.765
σ^2_1	0.355	0.466		
σ^2_2				45.915
σ^2_3			0.296	0.764
Goodness-of-fit				
Deviance	5133.137	5107.603	5110.798	5095.584
df	6	7	7	10
BIC	5172.217	5153.196	5156.390	5160.672

Model-based Trajectories, Employed vs Unemployed



Plotting Model-based Trajectories in R

```
tmp2 <- data.frame(x= seq(0, 15, by=5))
unemp_fe <- fixef(unemp_m4)

unemp= function(x){
       unemp_fe[1] + unemp_fe[2]*(1) + unemp_fe[3]*(1)*(x)
       }
emp1= function(x){
       ifelse(x >= 5,
       unemp_fe[1] + unemp_fe[2]*(0) + unemp_fe[3]*(0)*(x),
       NA)
       }
emp2= function(x){
       ifelse(x < 5,
       unemp_fe[1] + unemp_fe[2]*(0) + unemp_fe[3]*(0)*(x),
       NA)
       }
```

Plotting Model-based Trajectories in R

```
unemp_fit <- ggplot(data=tmp2, aes(x))

unemp_fit +
      stat_function(fun=unemp, size=1.5) +
        geom_text(aes(label="Unemployed", x=14.5, y= 14.75, size=1)) +
      stat_function(fun=emp1) +
      stat_function(fun=emp2, linetype="dashed") +
         geom_text(aes(label="Employed", x=14.5, y= 10.5, size=1)) +
      theme_bw() +
      theme(panel.grid.major = element_blank(),
            panel.grid.minor = element_blank()) +
      scale_x_continuous(limits=c(0, 15), "Months") +
      scale_y_continuous(limits=c(5, 20), "CES-D Score") +
      theme(legend.position="none")
```

Alternative coding strategies for TIME

WAVE	DAY	TIMEOFDAY	TIME	TIME-6.67	TIME-3.33
1	0	0.00	0.00	-6.67	-3.33
2	0	0.33	0.33	-6.34	-3.00
3	0	0.67	0.67	-6.00	-2.66
...					
16	5	0.00	5.00	-1.67	1.67
17	5	0.33	5.33	-1.34	2.00
18	6	0.00	6.00	-0.67	2.67
19	6	0.33	6.33	-0.34	3.00
20	6	0.67	6.67	0.00	3.34

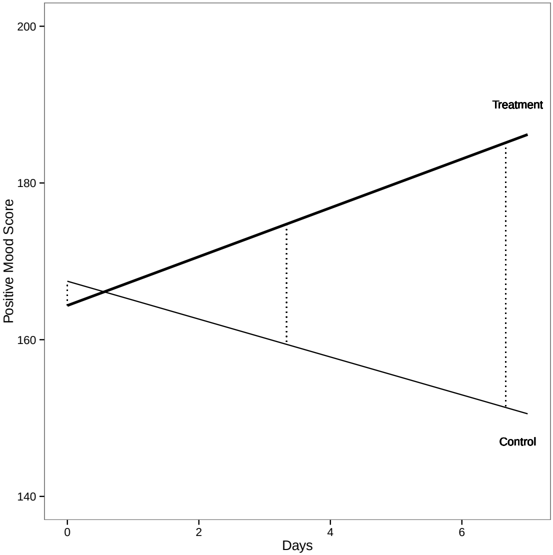
Estimating Alternative Coding Strategies for TIME in R

```
med_l <- read.dta("./data/med.dta", convert.factors=T)

med_l$time333 <- med_l$time-median(med_l$time)
med_l$time667 <- med_l$time-max(med_l$time)

## Initial Status
med_m1 <- lmer(pos ~ treat + time*treat + (time | id),
                 med_l,
                 REML = FALSE)

## Midpoint Status
med_m2 <- lmer(pos ~ treat + time333*treat + (time333 | id),
                 med_l,
                 REML = FALSE)

## Final Status
med_m3 <- lmer(pos ~ treat + time667*treat + (time667 | id),
                 med_l,
                 REML = FALSE)
```

Results of using alternative coding strategies for *TIME*

	TIME	TIME-3.33	TIME-6.67
Fixed Effects			
γ_{00}, Initial status	167.461	159.412	151.338
	(9.326)	(8.763)	(11.545)
γ_{01}, Treatment	−3.102	15.328	33.813
	(12.333)	(11.543)	(15.162)
γ_{10}, Rate of change	−2.417	−2.417	−2.417
	(1.731)	(1.731)	(1.731)
γ_{11}, Treatment	5.535	5.535	5.535
	(2.277)	(2.277)	(2.277)
Variance Components			
σ_ϵ^2	1229.992	1229.992	1229.992
σ_0^2	2111.555	2008.176	3323.874
σ_1^2	63.713	63.713	63.713
σ_{01}^2	−121.605	90.560	303.362
Goodness-of-fit			
Deviance	12680.51	12680.51	12680.51
df	8	8	8
BIC	12737.50	12737.50	12737.50

Plot for Alternative Coding Strategies for TIME in R



Plotting Alternative Coding Strategies for TIME in R

```
med_fe <- fixef(med_m1)
cont= function(x){
med_fe[1] + med_fe[2]*(0) + med_fe[3]*(x) + med_fe[4]*(0)*(x)
}
treat= function(x){
med_fe[1] + med_fe[2]*(1) + med_fe[3]*(x) + med_fe[4]*(1)*(x)
}

tmp1 <- data.frame(x= seq(min(med_l$time) : max(med_l$time)))


med_fit <- ggplot(data=tmp1, aes(x))
```

Plotting Alternative Coding Strategies for TIME in R

```
med_fit <- ggplot(data=tmp1, aes(x))

med_fit +
stat_function(fun=treat, size=1.1) +
 geom_text(aes(label="Treatment", x=6.85, y= 190, size=1)) +
stat_function(fun=cont) +
 geom_text(aes(label="Control", x=6.85, y= 147, size=1)) +
 geom_segment(aes(x= 0, y= 164.3585 , xend = 0, yend = 167.4608), linetyp
geom_segment(aes(x= 6.667, y= 185.1423, xend = 6.667, yend = 151.3457), l
geom_segment(aes(x= 3.333, y= 174.7489, xend = 3.333, yend = 159.4044), l
    theme_bw() +
    theme(panel.grid.major = element_blank(),
  panel.grid.minor = element_blank()) +
    xlab("Days") +
    scale_y_continuous(limits = c(140, 200),
                   "Positive Mood Score") +
theme(legend.position="none")
```

Polynomial Models, Example Data

- Data comes from Prosser, Rasbash and Goldstein (1991)
- 66 Asian children in a British community
- Four waves at approximately 6 weeks, and then 8, 12, and 27 months.
- 12% random sample stratified by gender

Time-varying Covariates, Example Data

- 4 waves of data–each subject has four records
- *AGE* in years
- *FEMALE*, $0 =$ Male to $2 =$ Female
- *WEIGHT* in kilograms

Final Model

$$Y_{ij} = \gamma_{00} + \gamma_{10} AGE_{ij} + \gamma_{20} AGE^2 + \zeta_{0i} + \zeta_{1i} AGE_{ij} + \epsilon_{ij}$$

$$\epsilon_{ij} \sim N(0, \sigma_\epsilon^2) \text{ and } \begin{bmatrix} \zeta_{0i} \\ \zeta_{1i} \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_0^2 & \sigma_{01} \\ \sigma_{10} & \sigma_2^2 \end{bmatrix} \right)$$

- γ_{00} is the *intercept*
- γ_{10} is the *instantaneous rate of change* when AGE is 0
- γ_{20} is the *curvature* or *de/acceleration* parameter associated with AGE^2.

Estimating Polynomial Models in R

```
wgt_m2 <- lmer(weight ~ age + I(age^2) + (age|id),
          data=wgt_l,
REML=F)

hlm.output(wgt_m2)
```

Model Results

	Model 1	Model 2	Model 3	Model 4	Model 5
Fixed Effects					
γ_{00}, Initial status	5.11	3.495	3.933	4.253	4.391
	(0.149)	(0.137)	(0.419)	(0.337)	(0.323)
γ_{01}, FEMALE			−0.288	−0.504	−0.596
			(0.267)	(0.207)	(0.196)
γ_{10}, AGE	3.46	7.704	8.886	8.045	7.698
	(0.126)	(0.239)	0.739	(0.352)	(0.238)
γ_{11}, AGE by FEMALE			−0.798	−0.230	
			(0.473)	(0.173)	
γ_{20}, AGE^2		−1.660	−1.993	−1.659	−1.658
		(0.089)	(0.273)	(0.088)	(0.088)
γ_{21}, AGE^2 by FEMALE			−0.225	−0.303	
			(0.175)	(0.112)	
Variance Components					
σ_ϵ^2	1.358	0.332	0.324	59.097	0.328
σ_0^2	0.097	0.404	0.356	45.765	0.354
σ_1^2	0.202	0.254	0.242		0.260
σ_{12}	0.140	0.088	0.054	0.764	0.048
Goodness-of-fit					
Deviance	685.4355	516.1557	504.3449	505.9897	507.7338
df	6	7	10	9	8.0000
BIC	717.1651	553.1736	557.2276	553.5841	550.0400

Final Model

$$Y_{ij} = \gamma_{00} + \gamma_{01} FEMALE_i + \gamma_{10} AGE_{ij} + \gamma_{20} AGE_{ij}^2 + \zeta_{0i} + \zeta_{1i} AGE_{ij} + \epsilon_{ij}$$

$$\epsilon_{ij} \sim N(0, \sigma_\epsilon^2) \text{ and } \begin{bmatrix} \zeta_{0i} \\ \zeta_{1i} \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_0^2 & \sigma_{01} \\ \sigma_{10} & \sigma_2^2 \end{bmatrix} \right)$$

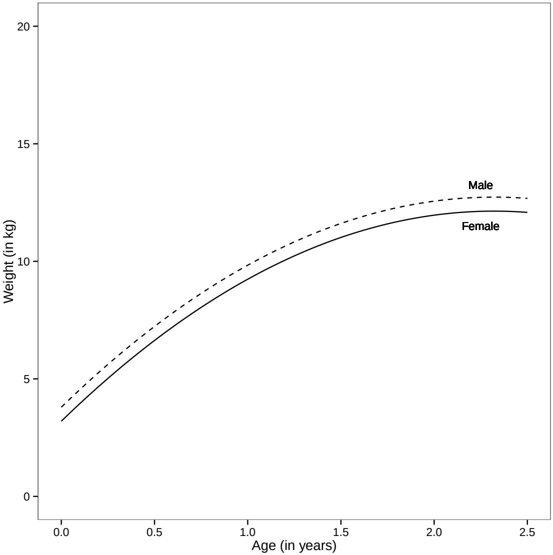
Estimating Polynomial Models in R

```
wgt_m5 <- lmer(weight ~ age + female + I(age^2) + (age|id),
            data=wgt_l,
            REML=F)

hlm.output(wgt_m5)
```

Model Results

	Model 1	Model 2	Model 3	Model 4	Model 5
Fixed Effects					
γ_{00}, Initial status	5.11	3.495	3.645	3.749	3.795
	(0.149)	(0.137)	(0.186)	(0.168)	(0.166)
γ_{01}, FEMALE			−0.288	−0.504	−0.596
			(0.267)	(0.207)	(0.196)
γ_{10}, AGE	3.46	7.704	8.089	7.815	7.698
	(0.126)	(0.239)	(0.328)	(0.253)	(0.238)
γ_{11}, AGE by FEMALE			−0.798	−0.230	
			(0.473)	(0.173)	
γ_{20}, AGE^2		−1.660	−1.767	−1.659	−1.658
		(0.089)	(0.121)	(0.088)	(0.088)
γ_{21}, AGE^2 by FEMALE			−0.225	−0.303	
			(0.175)	(0.112)	
Variance Components					
σ_ϵ^2	1.358	0.332	0.324	0.328	0.328
σ_0^2	0.097	0.404	0.356	0.347	0.354
σ_1^2	0.202	0.254	0.242	0.247	0.260
σ_{12}	0.140	0.088	0.054	0.055	0.048
Goodness-of-fit					
Deviance	685.4355	516.1557	504.3449	505.9897	507.7338
df	6	7	10	9	8.0000
BIC	717.1651	553.1736	557.2276	553.5841	550.0400

Plot for Quadratic Model Estimates



Plotting Polynomial Models in R

```
wgt_fe <- fixef(wgt_m5)

male= function(x){
   wgt_fe[1] + wgt_fe[2]*(x) + wgt_fe[3]*(0) + wgt_fe[4]*(x)^2
  }
female= function(x){
   wgt_fe[1] + wgt_fe[2]*(x) + wgt_fe[3]*(1) + wgt_fe[4]*(x)^2
  }

tmp <- data.frame(x= seq(0 , max(wgt_l$age), by=0.5 ))
```

Plotting Polynomial Models in in R

```
wgt_m5_fit <- ggplot(data=tmp, aes(x))

wgt_m5_fit +
  stat_function(fun=male, linetype="dashed") +
       geom_text(aes(label="Male", x=2.25, y=13.75, size=1)) +
  stat_function(fun=female) +
       geom_text(aes(label="Female", x=2.25, y=11.5, size=1)) +
  theme_bw() +
  theme(panel.grid.major = element_blank(),
        panel.grid.minor = element_blank()) +
  scale_x_continuous("Age (in years)") +
  scale_y_continuous(limits = c(0, 20),
                     "Weight (in kg)") +
  theme(legend.position="none")
```

The error covariance matrix

- Error covariance structures

A Generic Model

$$Y_{ij} = [\gamma_{00} + \gamma_{10} TIME_j + \gamma_{10}(X) + \gamma_{11}(X) TIME_j] + [\epsilon_{ij} + \zeta_{0i} + \zeta_{1i} TIME_j]$$

$$\begin{bmatrix} \zeta_{0i} \\ \zeta_{1i} \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_0^2 & \sigma_{01} \\ \sigma 10 & \sigma_1^2 \end{bmatrix} \right)$$

$$r_{ij} = [\epsilon_{ij} + \zeta_{0i} + \zeta_{1i} TIME_j]$$

$$Y_{ij} = [\gamma_{00} + \gamma_{10} TIME_j + \gamma_{10}(X) + \gamma_{11}(X) TIME_j] + r_{ij}$$

Diagonal with Homoscedastic Variance

$$\begin{bmatrix} r_{11} \\ r_{12} \\ r_{13} \\ r_{14} \\ r_{21} \\ r_{22} \\ r_{23} \\ r_{24} \\ \vdots \\ r_{n1} \\ r_{n2} \\ r_{n3} \\ r_{n4} \end{bmatrix} \sim N \left(\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} \sigma_r^2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 \\ 0 & \sigma_r^2 & 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 \\ 0 & 0 & \sigma_r^2 & 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \sigma_r^2 & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \sigma_r^2 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \sigma_r^2 & 0 & 0 & \dots & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \sigma_r^2 & 0 & \dots & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \sigma_r^2 & \dots & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots & \sigma_r^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 & \sigma_r^2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 & \sigma_r^2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \sigma_r^2 \end{bmatrix} \right)$$

Block Diagonal with Heteroscedastic Variance

$$\begin{bmatrix} r_{11} \\ r_{12} \\ r_{13} \\ r_{14} \\ r_{21} \\ r_{22} \\ r_{23} \\ r_{24} \\ \vdots \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ \vdots \end{bmatrix} \begin{bmatrix} \sigma^2_{r1} & \sigma_{r_1 r_2} & \sigma_{r_1 r_3} & \sigma_{r_1 r_4} & 0 & 0 & 0 & 0 & \dots \\ \sigma_{r_2 r_1} & \sigma^2_{r_2} & \sigma_{r_2 r_3} & \sigma_{r_2 r_4} & 0 & 0 & 0 & 0 & \dots \\ \sigma_{r_3 r_1} & \sigma_{r_3 r_2} & \sigma^2_{r_3} & \sigma_{r_3 r_4} & 0 & 0 & 0 & 0 & \dots \\ \sigma_{r_4 r_1} & \sigma_{r_4 r_2} & \sigma_{r_4 r_3} & \sigma^2_{r_4} & 0 & 0 & 0 & 0 & \dots \\ 0 & 0 & 0 & 0 & \sigma^2_{r1} & \sigma_{r_1 r_2} & \sigma_{r_1 r_3} & \sigma_{r_1 r_4} & \dots \\ 0 & 0 & 0 & 0 & \sigma_{r_2 r_1} & \sigma^2_{r_2} & \sigma_{r_2 r_3} & \sigma_{r_2 r_4} & \dots \\ 0 & 0 & 0 & 0 & \sigma_{r_3 r_1} & \sigma_{r_3 r_2} & \sigma^2_{r_3} & \sigma_{r_3 r_4} & \dots \\ 0 & 0 & 0 & 0 & \sigma_{r_4 r_1} & \sigma_{r_4 r_2} & \sigma_{r_4 r_3} & \sigma^2_{r_4} & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix} \right)$$

$$\mathbf{r} \sim N\left(\mathbf{0}, \begin{bmatrix} \Sigma_r & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \Sigma_r & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \Sigma_r & \dots & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \Sigma_r \end{bmatrix} \right)$$

Unstructured Error Covariance Matrix

$$\Sigma_r = \begin{bmatrix} \sigma^2_{r_1} & \sigma_{r_1 r_2} & \sigma_{r_3 r_1} & \sigma_{r_4 r_1} \\ \sigma_{r_2 r_1} & \sigma^2_{r_2} & \sigma_{r_3 r_2} & \sigma_{r_4 r_2} \\ \sigma_{r_3 r_1} & \sigma_{r_3 r_2} & \sigma^2_{r_3} & \sigma_{r_4 r_3} \\ \sigma_{r_4 r_1} & \sigma_{r_4 r_2} & \sigma_{r_3 r_4} & \sigma^2_{r_4} \end{bmatrix}$$

- Its deviance will always be the smallest of any error covariance structure.
- Will likely have a larger BIC due to “wasted” *df* from estimating all those parameters!

Autoregressive Error Covariance Matrix

$$\begin{bmatrix} \sigma^2 & \sigma^2\rho & \sigma^2\rho^2 & \sigma^2\rho^3 \\ \sigma^2\rho & \sigma^2 & \sigma^2\rho & \sigma^2\rho^2 \\ \sigma^2\rho^2 & \sigma^2\rho & \sigma^2 & \sigma^2\rho \\ \sigma^2\rho^3 & \sigma^2\rho^2 & \sigma^2\rho & \sigma^2 \end{bmatrix}$$

- "Band-diagonal" shape has an intuitive appeal for growth processes.
- Main diagonal (variances) are *homoscedastic.*
- Band-diagonal elements have identical covariances.
- ρ is the error autocorrelation parameter.
- Saves *df* because only two parameters are estimated.

Heterogeneous Autoregressive Error Covariance Matrix

$$\begin{bmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho & \sigma_1\sigma_3\rho^2 & \sigma_1\sigma_4\rho^3 \\ \sigma_2\sigma_1\rho & \sigma_2^2 & \sigma_2\sigma_3\rho & \sigma_2\sigma_4\rho^2 \\ \sigma_3\sigma_1\rho^2 & \sigma_3\sigma_2\rho & \sigma_3^2 & \sigma_3\sigma_4\rho \\ \sigma_4\sigma_1\rho^3 & \sigma_4\sigma_2\rho^2 & \sigma_4\sigma_3\rho & \sigma_4^2 \end{bmatrix}$$

- A relaxed version of the autoregressive structure–the main diagonal elements are *heteroscedastic*
- Band-diagonal elements are free to differ in magnitude.
- Uses more *df* than the strict autoregressive structure but is more efficient than the unstructured error covariance model

Toepliz Error Covariance Matrix

$$\begin{bmatrix} \sigma^2 & \sigma_1 & \sigma_2 & \sigma_3 \\ \sigma_1 & \sigma^2 & \sigma_1 & \sigma_2 \\ \sigma_2 & \sigma_1 & \sigma^2 & \sigma_1 \\ \sigma_3 & \sigma_2 & \sigma_1 & \sigma^2 \end{bmatrix}$$

- Band-diagonal elements have identical covariances.
- Covariances are not forced to be a fraction of the variances.
- Each band's magnitude is determined by the data.