

Still More Multilevel Models

2017-03-24

Preparation

We will use two datasets located in the *nbaLevel1.sav* file and the *nbaLevel2.sav* file. These data include player-level attributes for $n = 300$ NBA players, and team-level attributes for $N = 30$ different teams. The source of these data is: Woltman, Feldstein, MacKay, & Rocchi (2012). We will use these data to explore the question of whether how good a player is (Shots_on_five) predicts variation in life satisfaction.

The player-level attributes in the *nbaLevel1.sav* file include:

- Team_ID: The team ID number for each player
- Shots_on_five: A proxy for player quality/success. This indicates the number of successful shots (out of five taken). Higher values indicate a more successful player.
- Life_Satisfaction: Score on a survey of life satisfaction. Scores range from 5 to 25, with higher scores indicating more life satisfaction.

The team-level attributes in the *nbaLevel2.sav* file include:

- Team_ID: The team ID number
- Coach_Experience: Years of coaching experience in the NBA

We will use the `read_sav()` function from the **haven** library to import the data. We will also merge the two datasets together using the `left_join()` function from the **dplyr** package.

```
# Read in player-level and team-level data
library(haven)
nbaL1 = read_sav(file = "~/Google Drive/Documents/EPsy-8252/data/nbaLevel1.sav")
nbaL2 = read_sav(file = "~/Google Drive/Documents/EPsy-8252/data/nbaLevel2.sav")

# Merge the datasets
library(dplyr)
nba = left_join(nbaL1, nbaL2, by = "Team_ID")
head(nba)
```

Team_ID	Shots_on_five	Life_Satisfaction	Coach_Experience
01	3	18.8	2
01	3	18.0	2
01	4	21.0	2
01	4	20.5	2
01	3	19.0	2
01	2	12.1	2

```
# Load other libraries
library(AICcmodavg)
library(broom)
library(ggplot2)
library(lme4)
library(sm)
```

Estimation: Maximum Likelihood (ML) vs. Restricted Maximum Likelihood (REML)

Before we consider the model selection process, we need to discuss two common estimation methods for multilevel models: Maximum likelihood (ML) and Restricted maximum likelihood (REML). In conventional regression, we used ML estimation to estimate the coefficients and residual standard error. Remember that in ML estimation the estimates produced are the coefficient values and the value of the error variance that maximize the likelihood (or log-likelihood) function. For example, the log-likelihood for the intercept-only model in conventional regression is expressed as,

$$-\frac{N}{2} \times (2\pi\sigma_\epsilon^2) - \frac{1}{2\sigma_\epsilon^2} \times \sum \left(y_i - \beta_0 \right)^2$$

Using calculus, you can prove that the estimates for β_0 and σ_ϵ^2 that maximizes the log-likelihood are $\beta_0 = \frac{\sum y_i}{N}$ and $\sigma_\epsilon^2 = \frac{\sum \epsilon_i^2}{N}$, respectively. The estimate for the error variance

$$\sigma_\epsilon^2 = \frac{\sum \epsilon_i^2}{N}$$

is a biased estimate of the population error variance. It *underestimates* the population value in repeated sampling. In mathematical notation,

$$E(\hat{\sigma}_\epsilon^2) < \sigma_\epsilon^2$$

To adjust for this bias, a different denominator is used in the estimate, namely,

$$\text{Less biased } \sigma_\epsilon^2 = \frac{\sum \epsilon_i^2}{N - k}$$

where k is the number of regression coefficients being estimated. In the intercept-only model, $k = 1$ and the less biased estimate of the error variance is,

$$\sigma_\epsilon^2 = \frac{\sum \epsilon_i^2}{N - 1}$$

This is why we divide by $n - 1$ in the computation of standard deviation. To get this less biased estimate of error variance, we need to use a different estimation method. The method of estimation that produces this is referred to as *restricted maximum likelihood* (REML). The REML estimate of the coefficient β_0 is still $\frac{\sum y_i}{N}$. (REML is the default method of estimation for the `lmer()` function.) To use ML rather than REML, we include the argument `REML=FALSE` in the `lmer()` function.

```
# Estimate model using ML
lmer.0 = lmer(Life_Satisfaction ~ 1 + (1 | Team_ID), data = nba, REML = FALSE)
```

In general, comparing ML estimates with REML estimates, you find that the big difference is in the variance estimates and the standard errors. In general, the coefficient estimates for the fixed-effects will be quite similar. When the sample size (N) is very large, the ML and REML estimates for variance and standard errors will be pretty much the same (this is because $N \approx N - k$ for large N). For smaller sample sizes, the difference between the ML and REML estimates for variance and standard errors is dependent on the value of k (the number of fixed-effects in the model).

There are several schools of thought on which estimation method should be used in practice. One school of thought is that you should use ML estimation in the selection of fixed effects and REML estimation in

the selection of random effects. Another school of thought is to always use ML estimation, but to compare the final adopted model's estimates to the REML estimates. Large differences in the estimates would suggest problems with the robustness of the model (e.g., small sample size, overspecified model). (Using ML also allows for the use of common inferential methods such as deviance testing.) Still another school of thought is that one should always use REML estimation to allow for better variance estimate with low sample sizes.

In this set of notes, we will use ML estimation during the selection process when we are adopting the fixed-effects structure (i.e., choosing level-1 and level-2 predictors), and we will use REML estimation to select the random-effects structure (i.e., which random-effects to include). Once we have adopted a final model, we will use REML estimation to report and interpret any final set of presented models.

Model Building for Multilevel Models

The model selection process includes the following steps:

1. Fit the unconditional random intercepts model to identify the primary level of the unexplained variation. (Use REML)
2. Adopt the structural form for the level-1 model. This includes adding predictors (and potential interactions between predictors) to explain level-1 variation. In addition, you need to select the functional form of the predictors in the level-1 model (e.g., any polynomial or non-linear terms). (Use ML)
3. Adopt the structural form for the level-2 model(s). This includes adding predictors (and potential interactions between predictors) to explain level-2 variation. In addition, you need to select the functional form of the predictors in the level-2 model(s) (e.g., any polynomial or non-linear terms). (Use ML)
4. Adopt the random-effects structure for the level-2 model(s). (Use REML)

Sources of Variation: Fitting the Unconditional Random Intercepts Model

The first model that should be fitted in the multilevel model building process is the *unconditional random intercepts* model. This model is unconditional in that it includes no predictors (only an intercept). This intercept is allowed to vary across teams. Because the intercept-only model essentially estimates the mean, sometimes this model is referred to as the *unconditional means* model. The multilevel set of equations for this model are,

Level-1 :

$$\text{Life Satisfaction}_{ij} = \beta_{0j} + \epsilon_{ij}$$

Level-2 :

$$\beta_{0j} = \gamma_{00} + u_{0j}$$

And, the mixed-effects, or composite model is

$$\text{Life Satisfaction}_{ij} = \gamma_{00} + u_{0j} + \epsilon_{ij}$$

In these equations, γ_{00} , the fixed-effect, is the average life satisfaction score across teams. The u_{0j} term indicates the team-specific deviation in average life satisfaction for team j . Finally, ϵ_{ij} indicates the deviation between player i 's life satisfaction score and his team-specific average. We fit this model below.

```
lmer.0 = lmer(Life_Satisfaction ~ 1 + (1 | Team_ID), data = nba, REML = FALSE)
summary(lmer.0)
```

```
## Linear mixed model fit by maximum likelihood ['lmerMod']
## Formula: Life_Satisfaction ~ 1 + (1 | Team_ID)
## Data: nba
##
##      AIC      BIC   logLik deviance df.resid
##    1733     1744    -864     1727      297
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -2.7363 -0.8306  0.0705  0.8738  1.5604
##
## Random effects:
## Groups   Name      Variance Std.Dev.
## Team_ID (Intercept) 14.4      3.80
## Residual              14.6      3.82
## Number of obs: 300, groups: Team_ID, 30
##
## Fixed effects:
##              Estimate Std. Error t value
## (Intercept)   14.807      0.727    20.4
```

The estimated level-1 model,

$$\text{Life Satisfaction}_{ij} = 14.81,$$

indicating the predicted average life satisfaction score for all players (across teams) is 14.81.

The estimated variance for the random-effect of intercept provides a measure of the between-team (team-to-team) variation of life satisfaction scores. Looking at the variance components, we find that there seems to be between-team variation in life satisfaction ($\hat{\tau}_{00} = 14.96$) that we can explain by including team-level (level-2) predictors. The estimated residual variance provides a measure of the within-team (player-to-player) variation of life satisfaction scores. The non-zero residual variance ($\hat{\sigma}_{\epsilon}^2 = 14.61$) indicates including player-level (level-1) predictors may also be worthwhile in explaining variation in life satisfaction.

Intraclass Correlation Coefficient (ICC)

The ICC indicates the proportion of variance that is at level-2 (i.e., the proportion of unexplained variation that is between-team). To compute the ICC, we always use information from the unconditional random intercepts model;

$$\text{ICC} = \frac{\hat{\tau}_{00}}{\hat{\tau}_{00} + \hat{\sigma}_{\epsilon}^2}$$

For our example,

$$\text{ICC} = \frac{14.4}{14.4 + 14.6} = 0.497$$

Interpreting this value, an estimated 49.7% of the total unexplained variation in life satisfaction scores is attributable to differences between teams (at level-2). That implies that 50.3% of the total unexplained variation in life satisfaction scores is attributable to differences between players (at level-1).

Adopting the Level-1 Structure

Now that we know that including both level-1 and level-2 predictors is useful in explaining variation, we can start to adopt the fixed-effects structure of our model. The fixed-effects structure includes all non-random-effect components of both the level-1 and level-2 models. First, we will focus on adopting an appropriate level-1 model. (We have to initially adopt a level-1 model because the number of level-2 equations is dependent on the number of level-1 predictors!)

To begin, we will include potential player-level (level-1) predictors to the unconditional random intercepts model. These predictors will help explain the within-team variation. The only level-1 predictor we have in this dataset is player success (Shots_on_five). We will include this predictor in the level-1 model. Now that there are two level-1 coefficients (intercept and slope), we need two level-2 equations. The caveat being that the random-effects will be identical to the random-effects we specified in the unconditional intercepts model (namely ONLY a random-effect of intercept). The multilevel set of equations for this model are,

Level-1 :

$$\text{Life Satisfaction}_{ij} = \beta_{0j} + \beta_{1j}(\text{Player Success}) + \epsilon_{ij}$$

Level-2 :

$$\beta_{0j} = \gamma_{00} + u_{0j}$$

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

And, the mixed-effects, or composite model is

$$\text{Life Satisfaction}_{ij} = \gamma_{00} + \gamma_{10}(\text{Player Success}) + u_{0j} + \epsilon_{ij}$$

We fit this model using ML as,

```
# Estimate model using ML
lmer.1 = lmer(Life_Satisfaction ~ 1 + Shots_on_five + (1 | Team_ID), data = nba, REML = FALSE)
```

We also might consider whether there is a quadratic effect of player success. To do this, we will also fit a model that includes the quadratic effect of player success, again, keeping the same random-effects that we included in the unconditional random intercepts model. The multilevel set of equations for this model are,

Level-1 :

$$\text{Life Satisfaction}_{ij} = \beta_{0j} + \beta_{1j}(\text{Player Success}) + \beta_{2j}(\text{Player Success}^2) + \epsilon_{ij}$$

Level-2 :

$$\beta_{0j} = \gamma_{00} + u_{0j}$$

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

$$\beta_{2j} = \gamma_{20}$$

And, the mixed-effects, or composite model is

$$\text{Life Satisfaction}_{ij} = \gamma_{00} + \gamma_{10}(\text{Player Success}) + \gamma_{20}(\text{Player Success}^2) + u_{0j} + \epsilon_{ij}$$

We fit this model using ML as,

```
# Estimate model using ML
lmer.2 = lmer(Life_Satisfaction ~ 1 + Shots_on_five + I(Shots_on_five ^ 2) + (1 | Team_ID), data = nba, REML = FALSE)

# Compare the AICc values of the two models
# AICc evidence
aictab(
  cand.set = list(lmer.0, lmer.1, lmer.2),
  modnames = c("Unconditional model", "Linear Player success", "Quadratic Player success")
)
```

	Modnames	K	AICc	Delta_AICc	ModelLik	AICcWt	LL	Cum.Wt
3	Quadratic Player success	5	1385	0.000	1.000	0.563	-688	0.563
2	Linear Player success	4	1386	0.508	0.776	0.437	-689	1.000
1	Unconditional model	3	1733	347.970	0.000	0.000	-864	1.000

The evidence points toward the model that includes the linear and quadratic effects of player success as a level-1 predictors. There is also some evidence for the linear model. We will consider both models as we move forward. The next step in the process is to adopt the structural form for the level-2 models. To do this, we can include level-2 predictors in the level-2 equations. The only level-2 predictor we have is coach experience. This gives us three potential sets of level-2 equations:

- Coach experience should not be included in either level-2 equation (none).
- Coach experience should only be included in the level-2 equation for intercept (only intercept).
- Coach experience should be included in both level-2 equations (intercepts and slopes).

If we use the linear level-1 model, the multilevel models including coach experience in the level-2 equation for intercept only are:

Level-1 :

$$\text{Life Satisfaction}_{ij} = \beta_{0j} + \beta_{1j}(\text{Player Success}) + \epsilon_{ij}$$

Level-2 :

$$\beta_{0j} = \gamma_{00} + \gamma_{01}(\text{Coach Experience}) + u_{0j}$$

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

Again, using the linear level-1 model, the multilevel models including coach experience in both level-2 equation for intercept and slope are:

Level-1 :

$$\text{Life Satisfaction}_{ij} = \beta_{0j} + \beta_{1j}(\text{Player Success}) + \epsilon_{ij}$$

Level-2 :

$$\beta_{0j} = \gamma_{00} + \gamma_{01}(\text{Coach Experience}) + u_{0j}$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11}(\text{Coach Experience})$$

If we do not include coach experience as a predictor in any of the level-2 models, the multilevel models are the same as those we looked at when we were adopting the level-1 model, namely,

Level-1 :

$$\text{Life Satisfaction}_{ij} = \beta_{0j} + \beta_{1j}(\text{Player Success}) + \epsilon_{ij}$$

Level-2 :

$$\beta_{0j} = \gamma_{00} + u_{0j}$$

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

Putting these models in order from simplest to the most complex, and writing the composite (mixed-effects) model for each, we have,

$$1: \text{Life Satisfaction}_{ij} = \gamma_{00} + \gamma_{10}(\text{Player Success}_{ij}) + \left[u_{0j} + \epsilon_{ij} \right]$$

$$2: \text{Life Satisfaction}_{ij} = \gamma_{00} + \gamma_{10}(\text{Player Success}_{ij}) + \gamma_{01}(\text{Coach Experience}_{ij}) + \left[u_{0j} + \epsilon_{ij} \right]$$

$$3: \text{Life Satisfaction}_{ij} = \gamma_{00} + \gamma_{10}(\text{Player Success}_{ij}) + \gamma_{01}(\text{Coach Experience}_{ij}) + \gamma_{11}(\text{Coach Experience}_{ij})(\text{Player Success}_{ij}) + \left[u_{0j} + \epsilon_{ij} \right]$$

Note that including a level-2 predictor in the intercept equation produces a fixed-effect that is a main effect in the composite model. Including a level-2 predictor in the slope equation produces a fixed-effect that is an interaction. Recall that in order to interpret an interaction we also need the main effect to be included. Thus if we include a level-2 predictor in the slope equation, that predictor also has to be in the intercept equation.

Also note that again, the random-effects structure (only RE in the intercept) is kept the same as in the unconditional random intercepts model. In general, when comparing different structures of the fixed-effects, we need to keep the random-effects structure the same.

Below we fit these three models using `lmer()` with ML estimation.

```
# No level-2 predictors
lmer.l1 = lmer(Life_Satisfaction ~ 1 + Shots_on_five +
  (1 | Team_ID), data = nba, REML = FALSE)

# Level-2 predictor for intercept
lmer.l2 = lmer(Life_Satisfaction ~ 1 + Shots_on_five + Coach_Experience +
  (1 | Team_ID), data = nba, REML = FALSE)

# Level-2 predictor for intercept and slope
lmer.l3 = lmer(Life_Satisfaction ~ 1 + Shots_on_five + Coach_Experience +
  Shots_on_five:Coach_Experience + (1 | Team_ID), data = nba, REML = FALSE)
```

We were also considering the quadratic level-1 model. This gives us four potential sets of level-2 equations:

- Coach experience should not be included in any level-2 equation (none).
- Coach experience should only be included in the level-2 equation for intercept.
- Coach experience should be included in the level-2 equations for intercept and linear slope.
- Coach experience should be included in the level-2 equations for intercept, linear slope, and quadratic slope.

The four sets of multilevel models are:

Level-1 :

$$\text{Life Satisfaction}_{ij} = \beta_{0j} + \beta_{1j}(\text{Player Success}) + \beta_{2j}(\text{Player Success}^2) + \epsilon_{ij}$$

Level-2 :

$$\beta_{0j} = \gamma_{00} + u_{0j}$$

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

$$\beta_{2j} = \gamma_{20}$$

Level-1 :

$$\text{Life Satisfaction}_{ij} = \beta_{0j} + \beta_{1j}(\text{Player Success}) + \beta_{2j}(\text{Player Success}^2) + \epsilon_{ij}$$

Level-2 :

$$\beta_{0j} = \gamma_{00} + \gamma_{01}(\text{Coach Experience}) + u_{0j}$$

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

$$\beta_{2j} = \gamma_{20}$$

Level-1 :

$$\text{Life Satisfaction}_{ij} = \beta_{0j} + \beta_{1j}(\text{Player Success}) + \beta_{2j}(\text{Player Success}^2) + \epsilon_{ij}$$

Level-2 :

$$\beta_{0j} = \gamma_{00} + \gamma_{01}(\text{Coach Experience}) + u_{0j}$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11}(\text{Coach Experience})$$

$$\beta_{2j} = \gamma_{20}$$

Level-1 :

$$\text{Life Satisfaction}_{ij} = \beta_{0j} + \beta_{1j}(\text{Player Success}) + \beta_{2j}(\text{Player Success}^2) + \epsilon_{ij}$$

Level-2 :

$$\beta_{0j} = \gamma_{00} + \gamma_{01}(\text{Coach Experience}) + u_{0j}$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11}(\text{Coach Experience})$$

$$\beta_{2j} = \gamma_{20} + \gamma_{21}(\text{Coach Experience})$$

Again, we would write the composite models based on these multilevel models.

- 1: $\text{Life Satisfaction}_{ij} = \gamma_{00} + \gamma_{10}(\text{Player Success}_{ij}) + \gamma_{20}(\text{Player Success}_{ij}^2) + \left[u_{0j} + \epsilon_{ij} \right]$
- 2: $\text{Life Satisfaction}_{ij} = \gamma_{00} + \gamma_{10}(\text{Player Success}_{ij}) + \gamma_{20}(\text{Player Success}_{ij}^2) + \gamma_{01}(\text{Coach Experience}_{ij}) + \left[u_{0j} + \epsilon_{ij} \right]$
- 3: $\text{Life Satisfaction}_{ij} = \gamma_{00} + \gamma_{10}(\text{Player Success}_{ij}) + \gamma_{20}(\text{Player Success}_{ij}^2) + \gamma_{01}(\text{Coach Experience}_{ij}) + \gamma_{11}(\text{Coach Experience}_{ij})(\text{Player Success}_{ij}) + \left[u_{0j} + \epsilon_{ij} \right]$
- 4: $\text{Life Satisfaction}_{ij} = \gamma_{00} + \gamma_{10}(\text{Player Success}_{ij}) + \gamma_{20}(\text{Player Success}_{ij}^2) + \gamma_{01}(\text{Coach Experience}_{ij}) + \gamma_{11}(\text{Coach Experience}_{ij})(\text{Player Success}_{ij}) + \gamma_{21}(\text{Coach Experience}_{ij})(\text{Player Success}_{ij}^2) + \left[u_{0j} + \epsilon_{ij} \right]$

Below we fit these three models using `lmer()` with ML estimation.

```
# No level-2 predictors
lmer.q1 = lmer(Life_Satisfaction ~ 1 + Shots_on_five + I(Shots_on_five ^ 2) +
  (1 | Team_ID), data = nba, REML = FALSE)

# Level-2 predictor for intercept
lmer.q2 = lmer(Life_Satisfaction ~ 1 + Shots_on_five + I(Shots_on_five ^ 2) + Coach_Experience +
  (1 | Team_ID), data = nba, REML = FALSE)

# Level-2 predictor for intercept and slope
lmer.q3 = lmer(Life_Satisfaction ~ 1 + Shots_on_five + I(Shots_on_five ^ 2) + Coach_Experience +
  Shots_on_five:Coach_Experience + (1 | Team_ID), data = nba, REML = FALSE)

# Level-2 predictor for intercept, slope, and quadratic
lmer.q4 = lmer(Life_Satisfaction ~ 1 + Shots_on_five + I(Shots_on_five ^ 2) + Coach_Experience +
  Shots_on_five:Coach_Experience + I(Shots_on_five ^ 2):Coach_Experience +
  (1 | Team_ID), data = nba, REML = FALSE)
```

Now we will examine the model evidence for all seven models.

```
# AICc evidence
aictab(
  cand.set = list(lmer.l1, lmer.l2, lmer.l3, lmer.q1, lmer.q2, lmer.q3, lmer.q4),
  modnames = c("L1", "L2", "L3", "Q1", "Q2", "Q3", "Q4")
)
```

	Modnames	K	AICc	Delta_AICc	ModelLik	AICcWt	LL	Cum.Wt
7	Q4	8	1326	0.0	1	1	-655	1
5	Q2	6	1347	21.1	0	0	-667	1
2	L2	5	1347	21.1	0	0	-669	1
3	L3	6	1349	22.4	0	0	-668	1
6	Q3	7	1349	23.1	0	0	-667	1
4	Q1	5	1385	59.4	0	0	-688	1
1	L1	4	1386	59.9	0	0	-689	1

The evidence points to adoption of the level-1 model that includes both the linear and quadratic effects of player success, and the level-2 models that include coach experience as a predictor of intercept, linear effect, and quadratic effect. (If we were doing this in practice, we might also explore models that included a quadratic effect of coach experience in the level-2 models as well.)

Adopting a Random-Effects Structure

Now that we have a fixed-effect structure for the level-1 and level-2 models, we can explore which random effects structure to adopt. So far we have only explored having a random-effect of intercept. We can also explore models in which we allow a random-effect in the linear level-2 equation (in addition to the intercept), and one in which we allow a random-effect in the quadratic level-2 equation (in addition to the intercept and linear equations). We need to follow the hierarchy rule: If we allow a random-effect in a higher order equation (e.g., linear or quadratic), then a random-effect also needs to be included in all lower-order equations as well.

Let's explore the variance-covariance matrices associated with each of these random-effects structures. For the model that only included a random-effect of the intercept, the variance-covariance matrix was:

$$\begin{bmatrix} \tau_{00} \end{bmatrix}$$

Under this structure, we are estimating a single variance component, namely the variance associated with the team-level intercepts. How does this change if we allow for a random-effect of intercept and slope? The variance-covariance matrix for this random-effects structure is:

$$\begin{bmatrix} \tau_{00} & \tau_{01} \\ \tau_{10} & \tau_{11} \end{bmatrix}$$

Here we need to estimate three unique terms: (1) the variance of the team-specific intercepts (τ_{00}), (2) the variance of the team-specific linear slopes (τ_{11}), and (3) the covariance between the team-specific intercepts and slopes ($\tau_{01} = \tau_{10}$). How does this change if we allow for a random-effect of intercept, linear slope, and quadratic slope? The variance-covariance matrix for this random-effects structure is:

$$\begin{bmatrix} \tau_{00} & \tau_{01} & \tau_{02} \\ \tau_{10} & \tau_{11} & \tau_{12} \\ \tau_{20} & \tau_{21} & \tau_{22} \end{bmatrix}$$

Here we need to estimate six unique terms:

- The variance of the team-specific intercepts (τ_{00}).
- The variance of the team-specific linear slopes (τ_{11}).
- The variance of the team-specific quadratic slopes (τ_{22}).
- The covariance between the team-specific intercepts and linear slopes ($\tau_{01} = \tau_{10}$).
- The covariance between the team-specific intercepts and quadratic slopes ($\tau_{02} = \tau_{20}$).
- The covariance between the team-specific linear slopes and quadratic slopes ($\tau_{12} = \tau_{21}$).

Fitting each of these three random-effects structures is quite straight forward: we simply add the appropriate terms inside the parentheses where we specify the random-effects in the `lmer()` function. For the fixed-effects part of each model, we will use the model we adopted from the previous section. Lastly, since we are now estimating variance components, we will use REML estimation rather than ML to obtain better estimates.

```

# RE for intercept
lmer.re1 = lmer(Life_Satisfaction ~ 1 + Shots_on_five + I(Shots_on_five ^ 2) + Coach_Experience +
  Shots_on_five:Coach_Experience + I(Shots_on_five ^ 2):Coach_Experience +
  (1 | Team_ID), data = nba)

# RE for intercept and linear slope
lmer.re2 = lmer(Life_Satisfaction ~ 1 + Shots_on_five + I(Shots_on_five ^ 2) + Coach_Experience +
  Shots_on_five:Coach_Experience + I(Shots_on_five ^ 2):Coach_Experience +
  (1 + Shots_on_five | Team_ID), data = nba)

# RE for intercept, linear slope, and quadratic slope
lmer.re3 = lmer(Life_Satisfaction ~ 1 + Shots_on_five + I(Shots_on_five ^ 2) + Coach_Experience +
  Shots_on_five:Coach_Experience + I(Shots_on_five ^ 2):Coach_Experience +
  (1 + Shots_on_five + I(Shots_on_five ^ 2) | Team_ID), data = nba)

```

Then we obtain our model evidence,

```

# AICc evidence
aictab(
  cand.set = list(lmer.re1, lmer.re2, lmer.re3),
  modnames = c("RE1", "RE2", "RE3")
)

```

```

## Warning in aictab.AIClmerMod(cand.set = list(lmer.re1, lmer.re2, lmer.re3), :
## Model selection for fixed effects is only appropriate with ML estimation:
## REML (default) should only be used to select random effects for a constant set of fixed effects

```

Modnames	K	AICc	Delta_AICc	ModelLik	AICcWt	Res.LL	Cum.Wt
RE1	8	1341	0.00	1.000	0.833	-662	0.833
RE2	10	1344	3.29	0.193	0.160	-662	0.993
RE3	13	1350	9.66	0.008	0.007	-662	1.000

The model evidence suggests that we should adopt the random-effects structure having a random-effect for intercepts (but not for either slope term).

Model Reporting and Interpretation

Now that we have adopted a model, we should look at the REML output and report estimates and do any interpretation.

```
# Obtain REML estimates of the fixed-effects and variance components
summary(lmer.re1)

## Linear mixed model fit by REML ['lmerMod']
## Formula: Life_Satisfaction ~ 1 + Shots_on_five + I(Shots_on_five^2) +
##      Coach_Experience + Shots_on_five:Coach_Experience + I(Shots_on_five^2):Coach_Experience +
##      (1 | Team_ID)
##      Data: nba
##
## REML criterion at convergence: 1324
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -3.225 -0.681  0.026  0.653  2.784
##
## Random effects:
##      Groups      Name      Variance Std.Dev.
## Team_ID (Intercept) 0.0447  0.211
## Residual              4.6592  2.159
## Number of obs: 300, groups: Team_ID, 30
##
## Fixed effects:
##
##              Estimate Std. Error t value
## (Intercept)      8.1740     1.0264   7.96
## Shots_on_five    -1.8737     0.9491  -1.97
## I(Shots_on_five^2)  1.1078     0.2135   5.19
## Coach_Experience  -0.7038     0.6551  -1.07
## Shots_on_five:Coach_Experience  2.3764     0.4970   4.78
## I(Shots_on_five^2):Coach_Experience -0.4916     0.0958  -5.13
##
## Correlation of Fixed Effects:
##              (Intr) Shts__ I(Sh__^2) Cch_Ex S__:C_
## Shots_on_fv -0.824
## I(Shts__^2)  0.621 -0.931
## Coch_Exprnc -0.905  0.668 -0.456
## Shts_n_:C_E  0.851 -0.896  0.766  -0.868
## I(S__^2):C_ -0.712  0.921 -0.917   0.674 -0.933
```

The fixed-effects structure of the model includes an interaction term between the quadratic player success variable and coach experience. Because of this, we will not interpret any of the constituent effects. (E.g., we will not interpret either the linear or quadratic effect of player success, nor the main effect of coach experience.) Interpreting this interaction, we can say that there is a non-linear effect of player success on life satisfaction, and the non-linearity differs by level of coach experience. To better understand this effect, we should plot predicted life satisfaction against player success for different levels of coach experience.

Plotting the Fixed-Effects Part of the Model

We will do this in a similar way to how we plotted a conventional regression model. The difference is that in the `predict()` function, we need to specify that the prediction comes from only the fixed-effects part of the model. To do this, we include the argument `re.form = NA`. This tells the `predict()` function to ignore the random effects when making predictions. (It will give us the average models.)

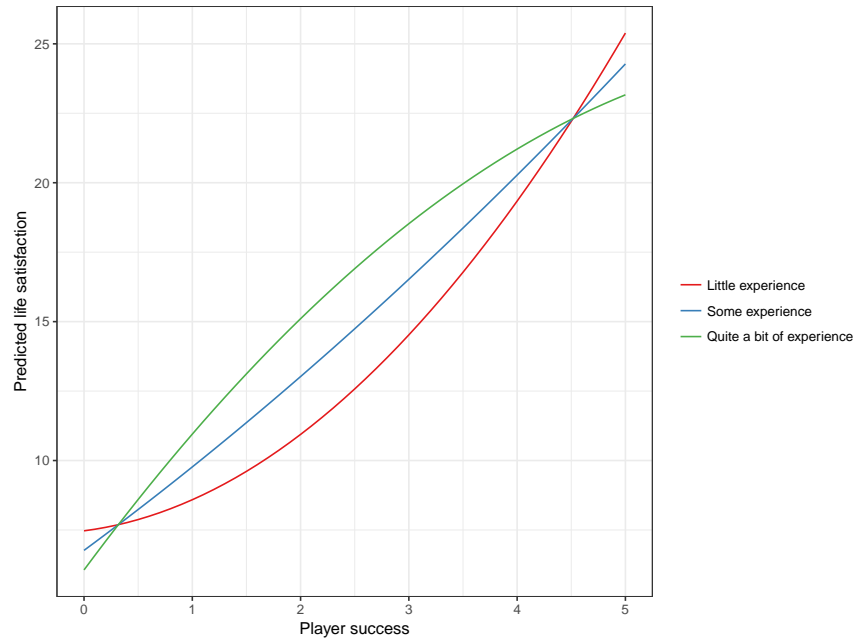
```
# Set up plotting data
plotdata = expand.grid(
  Shots_on_five = seq(from = 0, to = 5, by = 0.1),
  Coach_Experience = c(1, 2, 3)
)

# Predict life satisfaction
plotdata$yhat = predict(lmer.re1, newdata = plotdata, re.form = NA)
head(plotdata)
```

Shots_on_five	Coach_Experience	yhat
0.0	1	7.47
0.1	1	7.53
0.2	1	7.59
0.3	1	7.68
0.4	1	7.77
0.5	1	7.88

```
# Coerce coach experience into a factor for better plotting
plotdata$Coach_Experience = factor(plotdata$Coach_Experience, levels = c(1, 2, 3),
  labels = c("Little experience", "Some experience", "Quite a bit of experience"))

# Plot
ggplot(data = plotdata, aes(x = Shots_on_five, y = yhat, color = Coach_Experience)) +
  geom_line() +
  theme_bw() +
  xlab("Player success") +
  ylab("Predicted life satisfaction") +
  scale_color_brewer(name = "", palette = "Set1")
```



The plot suggests that for players on teams that have a coach with little experience, player success has a positive quadratic effect on life satisfaction. For players on teams that have a coach with some experience, player success has an almost linear (positive) effect on life satisfaction. And, for players on teams that have a coach with quite a bit of experience, player success has a negative quadratic effect on life satisfaction.

Taxonomy of Fitted Models

It is also common to present a selected taxonomy of fitted models in a table. You can of course decide which models to include in this table, but typically the “final” model from each of the fitting stages is presented. For us, this includes: (1) the unconditional random intercepts model, (2) the “final” level-1 model with quadratic effects of player success (and only random intercept), and (3) the “final” level-2 model with quadratic effects of player success and coach experience in all three level-2 equations (and only random intercept). (Note: If we had decided on a different RE structure, we would also include that model in the table.) Also despite having fitted some of these models using ML estimation, it is typical to present the results from REML for all the models presented in the table. This makes the effects from each model directly comparable to the others, since they are all fitted using the same estimation method.

	Model A	Model B	Model C
<i>Fixed-effects</i>			
Player success (L)		2.75 (0.39)	-1.87 (0.95)
Player success (Q)		0.12 (0.07)	1.11 (0.21)
Coach experience			-0.70 (0.66)
Player success (L) x Coach experience			2.38 (0.50)
Player success (q) x Coach experience			-0.49 (0.10)
Intercept	14.81 (0.74)	7.01 (0.50)	8.17 (1.03)
<i>Variance estimates</i>			
Team (Intercept)	14.96	0.91	0.04
Residual	14.61	5.24	4.66
<i>Goodness-of-Fit</i>			
AICc	1732	1392	1341

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$

Table 1. Coefficients (Standard Errors) for a Taxonomy of Fitted Multilevel (Player and Team) Models to Predict Life Satisfaction. All Models are Fitted using Restricted Maximum Likelihood.

Assumptions

Recall the assumptions we had for conventional regression: linearity, independence, normality of residuals, and homoscedasticity of residuals. We fitted the multilevel model to account for the non-independence of the residuals that arose because of clustering, so independence is not an assumption we need to examine here. The other assumptions—linearity, normality of residuals, and homoscedasticity of residuals—are also assumptions of the multilevel model (at least of the level-1 residuals), and need to be empirically evaluated. There is also an assumption that the distribution of each of the level-2 residuals is normally distributed.

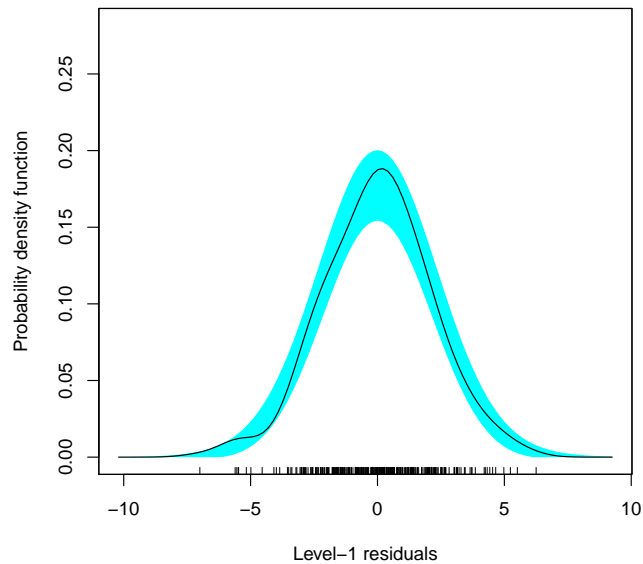
We can use the `augment()` function from the **broom** library to obtain the fitted values and the level-1 residuals.

```
out = augment(lmer.re1)
head(out)
```

Life_Satisfaction	Shots_on_five	I.Shots_on_five.2.	Coach_Experience	Team_ID	.fitted	.resid	.fixed	.mu	.
18.8	3	9	2	01	16.5	2.272	16.5	16.5	
18.0	3	9	2	01	16.5	1.468	16.5	16.5	
21.0	4	16	2	01	20.3	0.717	20.3	20.3	
20.5	4	16	2	01	20.3	0.217	20.3	20.3	
19.0	3	9	2	01	16.5	2.468	16.5	16.5	
12.1	2	4	2	01	13.0	-0.930	13.0	13.0	

We can examine the normality assumption of the level-1 residuals using the `sm.density()` function from the **sm** library.

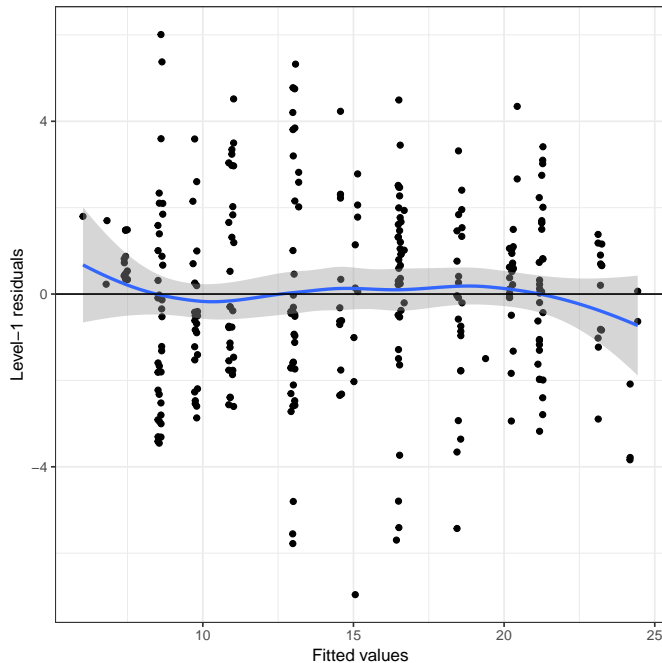
```
sm.density(out$.resid, model = "normal", xlab = "Level-1 residuals")
```



We also examine the other two assumptions by plotting the residuals versus the fitted values.

```
ggplot(data = out, aes(x = .fitted, y = .resid)) +  
  geom_point() +  
  geom_smooth() +  
  geom_hline(yintercept = 0) +  
  theme_bw() +  
  xlab("Fitted values") +  
  ylab("Level-1 residuals")
```

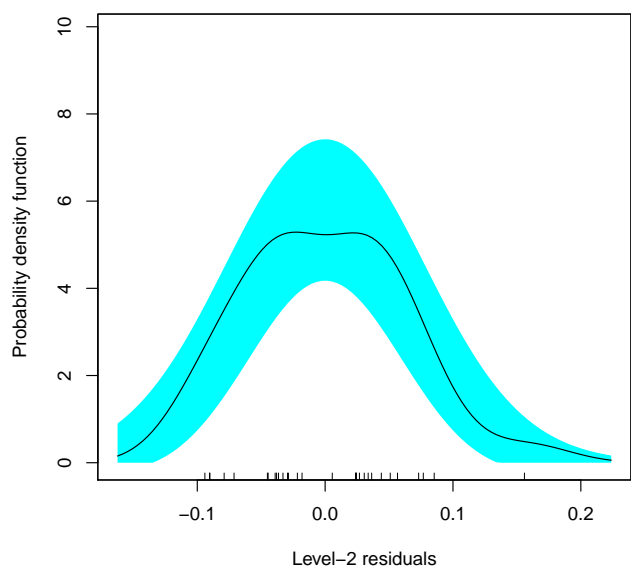
```
## 'geom_smooth()' using method = 'loess'
```

Examining the plots, it seems that the normality assumption seems satisfied. It also looks as though the linearity assumption appears to be satisfied. The homoscedasticity assumption is less clear. The range of residuals at fitted values near 15 seems greater than the range at fitted values less than 10 or greater than 20. That being said, the scale only ranges from -4 to 4 , so the greater range of the residuals around fitted values of 15 isn't that much bigger. (If we were concerned, we could try some sort of variance stabilizing transformation (e.g., log-transformation) on the outcome values. We would need to start the model selection process over if we transform the outcome.)

We will also examine the level-2 residuals to evaluate the normality assumption.

```
sm.density(ranef(lmer.re1)$Team_ID[, 1], model = "normal", xlab = "Level-2 residuals")
```



This plot suggests that the assumption the level-2 residuals are normally distributed may also be satisfied. (Note that if we had adopted a more complex random-effects structure, we would need to check this assumption for each set of level-2 residuals.)

References

Woltman, H., Feldstein, A., MacKay, J. C., & Rocchi, M. (2012). An introduction to hierarchical linear modeling. *Tutorials in Quantitative Methods for Psychology*, 8(1), 52–69.