

Mixed-Effects Models in Practice

2018-02-25

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

```
# Load libraries
library(AICcmodavg)
library(broom)
library(dplyr)
library(ggplot2)
library(haven) #for reading in the .sav files
library(lme4) #for fitting mixed-effects models
library(sm)

# Read in player-level data
nbaL1 = read_sav(file = "~/Dropbox/epsy-8252/data/nbaLevel1.sav")

# Read in team-level data
nbaL2 = read_sav(file = "~/Dropbox/epsy-8252/data/nbaLevel2.sav")

# Join/merge the two files together
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

Unconditional Means Model

As in a conventional fixed-effects regression analysis we begin a mixed-effects analysis by fitting the intercept-only model. This model is referred to as the *unconditional means model*. This model will include a fixed-effect of intercept and a random-effect of intercept, and no other predictors. This is the simplest model we can fit while still accounting for the dependence in the data (e.g., including a random-effect). The statistical model for this can be expressed as:

$$\text{Life Satisfaction}_{ij} = \beta_0 + b_{0j} + \epsilon_{ij},$$

where β_0 is the fixed-effect of intercept, b_{0j} is the random-effect of intercept for team j and ϵ_{ij} is the error for player i on team j . As we have been talking about in class, the full specification of a model also includes a mathematical description of the distributional assumptions. Mixed-effects models have distributional assumptions on the errors (ϵ_{ij}) and on each set of random-effects included in the model (b_{0j} in our model). The assumptions on the errors are:

- Independence;
- Conditional normality;
- Conditional means are 0; and
- Homoskedasticity of the conditional variances σ_ϵ^2 .

Note that the independence assumption does not assume independence in the original data, but is on the errors which are produced after we account for the dependence in the data by including a random-effect in the model.

The assumptions on each set of random effects are:

- Normality;
- Mean of 0; and
- There is some variance, $\sigma_{b_0}^2$ (often just denoted σ_0^2)

In mathematical notation the assumptions for the unconditional means model can be written as:

$$\epsilon_{ij} \stackrel{i.i.d}{\sim} \mathcal{N}(0, \sigma_\epsilon^2)$$

$$b_{0j} \sim \mathcal{N}(0, \sigma_0^2)$$

Fitting and Interpreting the Model

We fit this model and display the output below. We include the argument `REML=FALSE` to force the `lmer()` function to produce maximum likelihood estimates (rather than restricted maximum likelihood estimates). In practice, we will generally want to fit these models using ML estimation.

```
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 `summary()` function displays the fitted coefficients for the fixed-effects, and the variance estimates for the errors ($\hat{\sigma}_\epsilon^2$) and the random-effect of intercept ($\hat{\sigma}_0^2$). Using the fixed-effects estimates, the fitted equation for the fixed-effects model is:

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

We interpret coefficients from the fixed-effects model the same way we interpret coefficients produced from the `lm()` output. For example,

- The predicted average life satisfaction for all players is 14.81.

The variance estimates are:

- $\hat{\sigma}_\epsilon^2 = 14.6$
- $\hat{\sigma}_0^2 = 14.4$

Partitioning Variation

Let's re-look at the linear equation from the statistical model.

$$\text{Life Satisfaction}_{ij} = \beta_0 + b_{0j} + \epsilon_{ij}$$

In this model, the random-effect represents the deviation from Team j 's average life satisfaction score from the global average life satisfaction score. The error term represents the deviation between Player i 's life satisfaction score and his team's average life satisfaction score.

Let's consider the intercept-only fixed-effects regression model (no random-effect term):

$$\text{Life Satisfaction}_{ij} = \beta_0 + \epsilon_{ij}$$

In this model ALL the deviation from the average life satisfaction is at the player-level. When we fit a mixed-effects model,

$$\text{Life Satisfaction}_{ij} = \beta_0 + [b_{0j} + \epsilon_{ij}],$$

we can think of this model as partitioning that deviation from the average life satisfaction score into two parts: (1) the deviation to the team average, and (2) the player deviation to the team average.

In the same manner, the variance estimates have partitioned the variance into team-level variance and player-level variance. (Some statisticians may refer to these as *between-team* variance and *within-team* variance, respectively.) These represent variation that is UNACCOUNTED for by the model (they are errors/deviations after all).

One nice mathematical property of variances is that they are additive. Thus, we can compute the total unaccounted for variation by summing the team-level and player-level variance estimates (these are the only sources of variation in the model):

$$\begin{aligned}\hat{\sigma}_0^2 + \hat{\sigma}_\epsilon^2 \\ 14.4 + 14.6 = 29\end{aligned}$$

We can now use this value to compute the proportion of unaccounted for variation at both the team- and player-levels. The unaccounted for variation at the team-level is:

$$\frac{14.4}{29} = 0.497$$

The unaccounted for variation at the player-level is:

$$\frac{14.6}{29} = 0.503$$

Interpreting these,

- 49.7% of the unaccounted variation is team-level or between-team variation.
- 50.3% of the unaccounted variation is player-level or within-team variation.

This tells us that we could include both player-level predictors (to explain the unaccounted for player-level variation) and team-level predictors (to explain the unaccounted for team-level variation).

This partitioning of variation should be done in every analysis, and ALWAYS is done using the unconditional means model. The unconditional means model will serve as our *baseline* model. As we add predictors, we can compare the unaccounted for variation at each level in the predictor models to the baseline unaccounted for variation in the unconditional means model. This is one way of measuring how effective predictors are at further explaining variation in the model.

Including Player-Level Predictors

As we begin to include predictors in the model, we will begin with the player-level predictors. In the hierarchy of players and teams, players are at the lowest level (players are clustered within teams); as such, we refer to the player-level predictors as Level-1 predictors. (Team-level predictors in this model would be Level-2 predictors.) In a mixed-effects analysis, we start with the lowest level predictors and then work our way up.

In our data, shots-on-five is the only player-level (level-1) predictor. We begin by including the fixed-effect of shots-on-five in the model (with no random-effect for this predictor). The statistical model for this can be expressed as:

$$\text{Life Satisfaction}_{ij} = [\beta_0 + b_{0j}] + \beta_1(\text{SO5}) + \epsilon_{ij}$$

In this model, β_0 is the fixed-effect of intercept, b_{0j} is the random-effect of intercept for team j , β_1 is the fixed-effect of shots-on-five, and ϵ_{ij} is the error for player i on team j .

Fitting this using the `lmer()` function:

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

```
## Linear mixed model fit by maximum likelihood ['lmerMod']
## Formula: Life_Satisfaction ~ 1 + Shots_on_five + (1 | Team_ID)
## Data: nba
##
##      AIC      BIC   logLik deviance df.resid
##    1386    1401    -689    1378     296
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -2.4961 -0.7496  0.0496  0.6820  2.3318
##
## Random effects:
## Groups   Name      Variance Std.Dev.
## Team_ID  (Intercept) 0.812    0.901
## Residual                5.268    2.295
## Number of obs: 300, groups: Team_ID, 30
##
## Fixed effects:
##              Estimate Std. Error t value
## (Intercept)      6.414      0.364    17.6
## Shots_on_five    3.371      0.119    28.4
##
## Correlation of Fixed Effects:
##              (Intr)
## Shots_on_fv -0.814
```

Using the fixed-effects estimates, the fitted equation for the fixed-effects model is:

$$\text{Life Satisfaction}_{ij} = 6.41 + 3.37(\text{SO5})$$

Interpreting these coefficients,

- The predicted average life satisfaction for all players who make 0 of their five shots successfully is 6.41.
- On average, each additional successful shot is associated with a 3.37 increase in life satisfaction.

The variance estimates are:

- $\hat{\sigma}_{\epsilon}^2 = 5.27$
- $\hat{\sigma}_0^2 = 0.81$

First note that by including a player-level predictor we REDUCED the unaccounted for variation at the player-level. To determine how much we reduced it, we compute the proportion of reduction relative to the baseline model.

$$\frac{14.6 - 5.27}{14.6} = 0.639$$

Shots_on_five accounted for 63.9% of the unaccounted for variation at the player-level (level-1).

Including the player-level predictor also CHANGED the amount of unaccounted for variation at the team-level. In this case it happened to reduce this variation, but other times, you will see that the variation stays about the same, or increases! This is a mathematical artifact of the estimation. In a more practical sense, we wouldn't really be too interested in the team-level (level-2) variation at this point. We are only adding player-level (level-1) predictors to the model, so that is the variation that we expect to impact.

Including the Random-Effect of Shots-on-Five

We can also include a random-effect for any of our level-1 predictors. The statistical model for this can be expressed as:

$$\text{Life Satisfaction}_{ij} = [\beta_0 + b_{0j}] + [\beta_1(\text{SO5}) + b_{1j}(\text{SO5})] + \epsilon_{ij}.$$

In this model, β_0 is the fixed-effect of intercept, b_{0j} is the random-effect of intercept for team j , β_1 is the fixed-effect of shots-on-five, b_{1j} is the random-effect of shots-on-five for team j , and ϵ_{ij} is the error for player i on team j .

Note that the random-effect of shots-on-five is also multiplied by the predictor; same as the fixed-effect of shots-on-five. There are now two random-effects in the model; each indicating a team-level deviation to the respective fixed-effect.

Fitting this using the `lmer()` function:

```
lmer.2 = lmer(Life_Satisfaction ~ 1 + Shots_on_five + (1 + Shots_on_five | Team_ID),
             data = nba, REML = FALSE)
summary(lmer.2)
```

```
## Linear mixed model fit by maximum likelihood ['lmerMod']
## Formula: Life_Satisfaction ~ 1 + Shots_on_five + (1 + Shots_on_five |
##      Team_ID)
##      Data: nba
##
##      AIC      BIC    logLik deviance df.resid
##      1388     1410     -688    1376     294
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -2.5335 -0.6817  0.0241  0.7132  2.3847
##
## Random effects:
##      Groups   Name                Variance Std.Dev. Corr
##      Team_ID  (Intercept)    0.0978     0.313
##              Shots_on_five 0.0824     0.287    1.00
##      Residual                    5.1132     2.261
## Number of obs: 300, groups: Team_ID, 30
##
## Fixed effects:
##              Estimate Std. Error t value
## (Intercept)      6.397      0.316    20.2
## Shots_on_five     3.309      0.131    25.2
##
## Correlation of Fixed Effects:
##              (Intr)
## Shots_on_fv -0.742
```

Using the fixed-effects estimates, the fitted equation for the fixed-effects model is:

$$\widehat{\text{Life Satisfaction}}_{ij} = 6.41 + 3.37(\text{SO5})$$

Interpreting these coefficients,

- The predicted average life satisfaction for all players who make 0 of their five shots successfully is 6.40.
- On average, each additional successful shot is associated with a 3.31 increase in life satisfaction.

These estimates have not changed much from the previous model. There are now three sources of variation (team intercepts, team slopes, and random error), and thus three variance estimates. The variance estimates are:

- $\hat{\sigma}_{\epsilon}^2 = 5.11$
- $\hat{\sigma}_0^2 = 0.10$
- $\hat{\sigma}_1^2 = 0.08$

The player-level unaccounted for variance has not changed much (since we didn't add any player-level components into the model), but the team-level variance estimates have changed.

Choosing Random-Effects

You should always include a random-effect of intercept; this accounts for the dependence in the data. How can you decide which other play-level effects should include random-effects and which should be fixed? This is not always an easy decision.

One way to decide this is to explore the data visually. One plot that can be especially useful for this is a plot of the OLS regression lines plotted for each team. You can do this by plotting all the team lines in a single panel. It is also useful to show the global (average) line in this plot.

```
# Plot all the lines in a single panel
ggplot(data = nba, aes(x = Shots_on_five, y = Life_Satisfaction)) +
  geom_line(aes(group = Team_ID), stat = "smooth", method = "lm", se = FALSE,
            color = "blue", alpha = 0.7) +
  geom_line(aes(group = 1), stat = "smooth", method = "lm", se = FALSE,
            color = "black", size = 2) +
  theme_bw()
```

In this plot we can see that (1) the intercepts for each of the 30 team-level lines (blue) are not the same; and (2) the slopes for each of the 30 team-level lines are not the same. This suggests that we might consider a random-effect of intercept and a random-effect for the shots-on-five effect.

A second plot that is useful to examine is a plot of each team line displayed in a different panel. In this plot, it can also be useful to plot the players' data.

```
# Plot each team line in a different panel (not displayed)
ggplot(data = nba, aes(x = Shots_on_five, y = Life_Satisfaction)) +
  geom_point() +
  geom_smooth(aes(group = Team_ID), method = "lm", se = FALSE,
              color = "blue") +
  theme_bw() +
  facet_wrap(~Team_ID)
```

We can also see that the intercepts and slopes vary by team in this plot. Additionally, we can examine the functional form of the relationship by team. Here, it seems that the relationship between life satisfaction and shots-on-five seems reasonably linear for most of the teams.

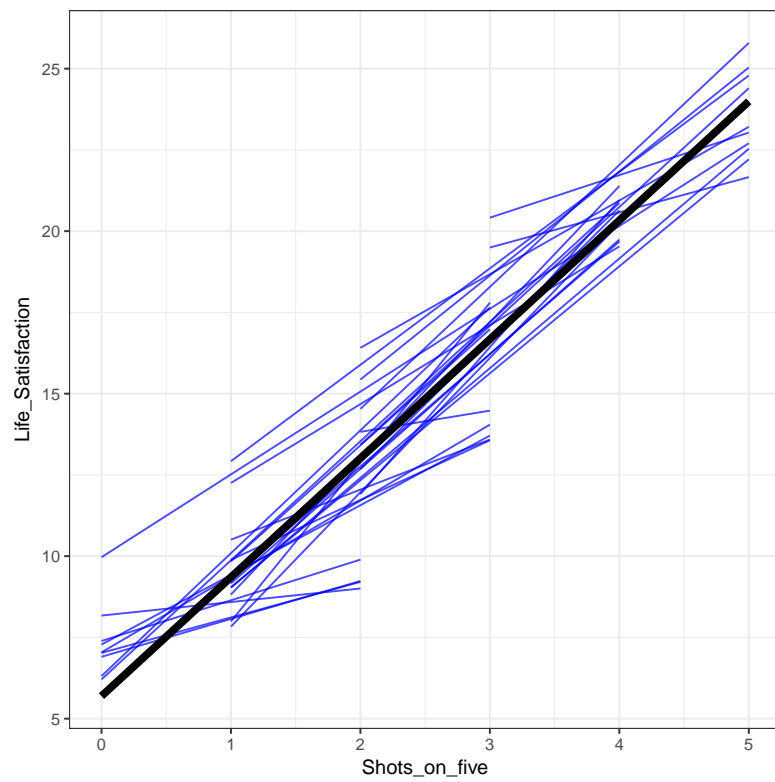


Figure 1: Plot of the OLS regression lines for each of the 30 NBA teams (blue). The OLS regression line for the 300 players ignoring team.

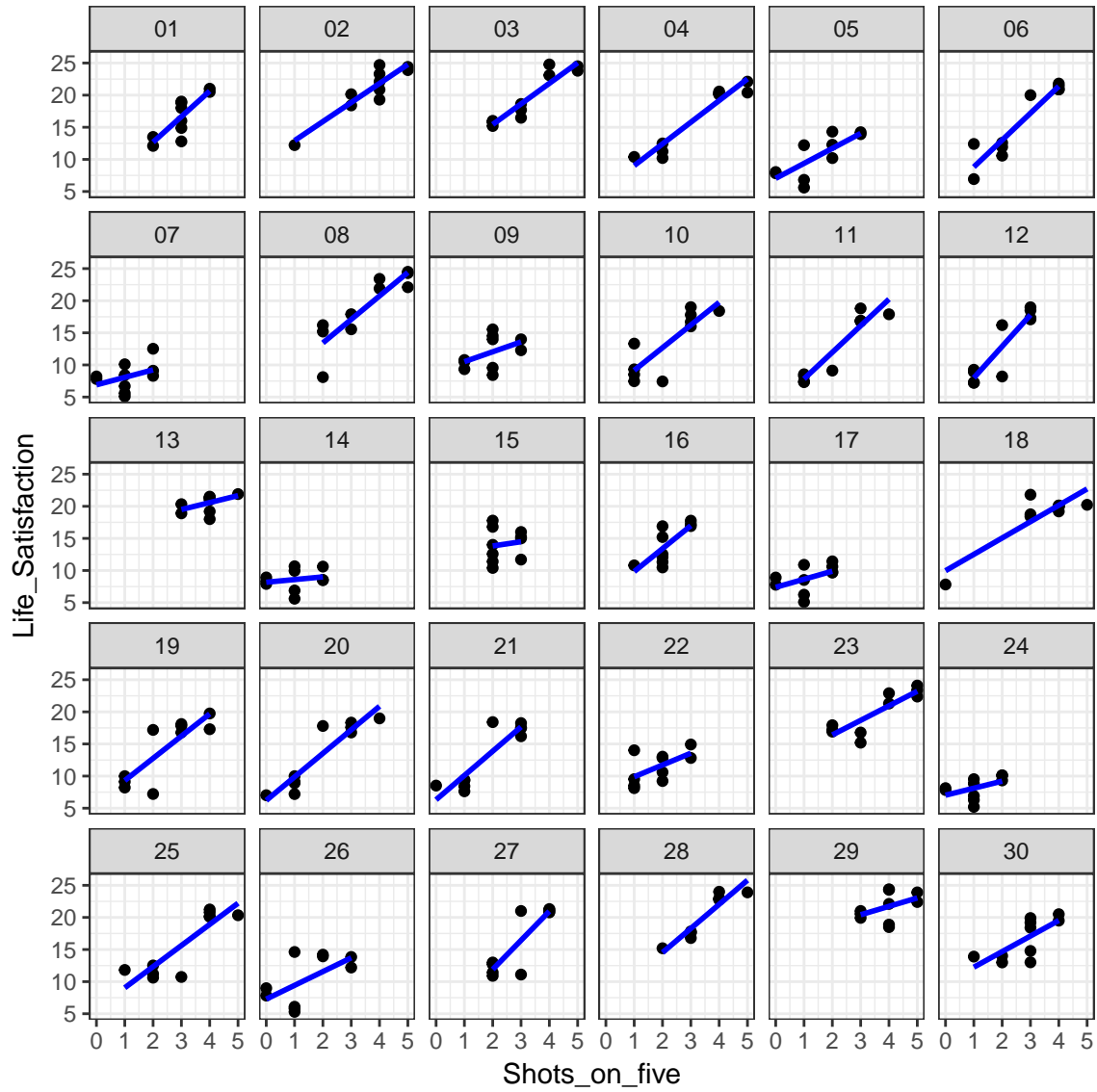


Figure 2: Scatterplot of the shots-on-five and life satisfaction data, as well as the OLS regression line for each of the 30 NBA teams.

Model Evidence

Since all of our models used the same data set and same outcome, and also were fitted using maximum likelihood, we can create a table of model evidence to help evaluate which random-effects are needed.

```
aictab(
  cand.set = list(lmer.0, lmer.1, lmer.2),
  modnames = c("B0 and b0", "B0, B1, and b0", "B0, B1, b0, and b1")
)
```

	Modnames	K	AICc	Delta_AICc	ModelLik	AICcWt	LL	Cum.Wt
2	B0, B1, and b0	4	1386	0.00	1.00	0.719	-689	0.719
3	B0, B1, b0, and b1	6	1388	1.88	0.39	0.281	-688	1.000
1	B0 and b0	3	1733	347.46	0.00	0.000	-864	1.000

Given the data and candidate models, the empirical evidence mostly supports including a fixed-effect of intercept and shots-on-five, and a random-effect of intercept (model probability = 0.72). There is some empirical support for also including a random-effect for slope. Before we commit, we will also explore adding in team-level predictors.

Including Team-Level Predictors

In our data, coach experience is the only team-level (level-2) predictor. We can include level-2 predictors in the model as fixed-effects. We DO NOT include random-effects for the level-2 predictors when there are only two levels of hierarchy. Since the table of model evidence indicated that there were two plausible random-effects structures, we will include coaching experience in the two models that have the most empirical evidence and compare them. The two statistical models can be expressed as:

$$\mathbf{M3} : \text{Life Satisfaction}_{ij} = [\beta_0 + b_{0j}] + \beta_1(\text{SO5}) + \beta_2(\text{Coach Experience}) + \epsilon_{ij}$$

$$\mathbf{M4} : \text{Life Satisfaction}_{ij} = [\beta_0 + b_{0j}] + [\beta_1(\text{SO5}) + b_1(\text{SO5})] + \beta_2(\text{Coach Experience}) + \epsilon_{ij}$$

We fit these models using the `lmer()` function as:

```
lmer.3 = lmer(Life_Satisfaction ~ 1 + Shots_on_five + Coach_Experience + (1 | Team_ID),
  data = nba, REML = FALSE)
lmer.4 = lmer(Life_Satisfaction ~ 1 + Shots_on_five + Coach_Experience +
  (1 + Shots_on_five | Team_ID), data = nba, REML = FALSE)
```

We can now examine the table of model evidence for all five candidate models (including the three previous models we fitted):

```
aictab(
  cand.set = list(lmer.0, lmer.1, lmer.2, lmer.3, lmer.4),
  modnames = c("B0 and b0", "B0, B1, and b0", "B0, B1, b0, and b1",
    "B0, B1, B2, and b0", "B0, B1, B2, b0, and b1")
)
```

	Modnames	K	AICc	Delta_AICc	ModelLik	AICcWt	LL	Cum.Wt
4	B0, B1, B2, and b0	5	1347	0.0	1.000	0.81	-669	0.81
5	B0, B1, B2, b0, and b1	7	1350	2.9	0.234	0.19	-668	1.00
2	B0, B1, and b0	4	1386	38.8	0.000	0.00	-689	1.00

	Modnames	K	AICc	Delta_AICc	ModelLik	AICcWt	LL	Cum.Wt
3	B0, B1, b0, and b1	6	1388	40.6	0.000	0.00	-688	1.00
1	B0 and b0	3	1733	386.2	0.000	0.00	-864	1.00

The majority of the empirical evidence supports the model that includes (1) fixed-effects of intercept, shots-on-five, and coach experience; and (2) a random-effect of intercept. (There is still some evidence supporting the inclusion of the shots-on-five random-effect.) Below we display the fitted coefficients and variance components.

```
summary(lmer.3)
```

```
## Linear mixed model fit by maximum likelihood ['lmerMod']
## Formula: Life_Satisfaction ~ 1 + Shots_on_five + Coach_Experience + (1 |
## Team_ID)
## Data: nba
##
##      AIC      BIC    logLik deviance df.resid
##    1347    1366     -668    1337     295
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -3.0161 -0.7515  0.0098  0.7919  2.6641
##
## Random effects:
## Groups   Name      Variance Std.Dev.
## Team_ID  (Intercept) 0.00     0.00
## Residual                5.05     2.25
## Number of obs: 300, groups: Team_ID, 30
##
## Fixed effects:
##              Estimate Std. Error t value
## (Intercept)      3.947      0.368   10.74
## Shots_on_five      3.134      0.122   25.60
## Coach_Experience    1.553      0.214    7.27
##
## Correlation of Fixed Effects:
##              (Intr) Shts__
## Shots_on_fv -0.154
## Coch_Exprnc -0.654 -0.591
```

Using the fixed-effects estimates, the fitted equation for the fixed-effects model is:

$$\text{Life Satisfaction}_{ij} = 6.41 + 3.37(\text{SO5} + 1.55(\text{Coach Experience}))$$

Interpreting these coefficients,

- The predicted average life satisfaction for all players who successfully make 0 of their five shots, and whose coach has 0 years of experience is 3.95.
- Controlling for differences in coach experience, each additional successful shot is associated with a 3.13 increase in life satisfaction, on average.
- Controlling for differences in number of shots made, each additional year of coach experience is associated with a 1.55 increase in life satisfaction, on average.

In this model, there are two sources of variation (team intercepts, and random error), and thus two variance estimates. The variance estimates are:

- $\hat{\sigma}_\epsilon^2 = 5.05$
- $\hat{\sigma}_0^2 < .01$

Cross-Level Interaction

Lastly, we will fit the model that includes an interaction between shots-on-five and coach experience. Because these predictors are at different levels this is sometimes referred to as a cross-level interaction. We use the same colon (:) notation to fit interactions in mixed-effects models as we did in fixed-effects models. For completeness, we fit this interaction in a model that includes the random-effect of intercept, and in a model that includes both the random-effect of intercept and shots-on-five.

```
lmer.5 = lmer(Life_Satisfaction ~ 1 + Shots_on_five + Coach_Experience +
             Shots_on_five:Coach_Experience + (1 | Team_ID),
             data = nba, REML = FALSE)
lmer.6 = lmer(Life_Satisfaction ~ 1 + Shots_on_five + Coach_Experience +
             Shots_on_five:Coach_Experience + (1 + Shots_on_five | Team_ID),
             data = nba, REML = FALSE)
```

Note that we do not include random-effects of the interaction term. We can now examine the table of model evidence for all seven candidate models fitted:

```
aictab(
  cand.set = list(lmer.0, lmer.1, lmer.2, lmer.3, lmer.4, lmer.5, lmer.6),
  modnames = c("B0 and b0", "B0, B1, and b0", "B0, B1, b0, and b1",
               "B0, B1, B2, and b0", "B0, B1, B2, b0, and b1",
               "B0, B1, B2, B1:B2, and b0",
               "B0, B1, B2, B1:B2, b0, and b1")
)
```

	Modnames	K	AICc	Delta_AICc	ModelLik	AICcWt	LL	Cum.Wt
4	B0, B1, B2, and b0	5	1347	0.00	1.000	0.526	-669	0.526
6	B0, B1, B2, B1:B2, and b0	6	1349	1.33	0.514	0.271	-668	0.797
5	B0, B1, B2, b0, and b1	7	1350	2.90	0.234	0.123	-668	0.920
7	B0, B1, B2, B1:B2, b0, and b1	8	1351	3.78	0.151	0.080	-667	1.000
2	B0, B1, and b0	4	1386	38.77	0.000	0.000	-689	1.000
3	B0, B1, b0, and b1	6	1388	40.65	0.000	0.000	-688	1.000
1	B0 and b0	3	1733	386.23	0.000	0.000	-864	1.000

The empirical evidence here is a little more split. Our top two models, which both show empirical support include the random-effect of intercept. There is also, some fairly compelling evidence supporting the model that includes an interaction between shots-on-five and coach experience and a random-effect of intercept. We will display and interpret the effects from this model below.

```
summary(lmer.5)

## Linear mixed model fit by maximum likelihood ['lmerMod']
## Formula:
## Life_Satisfaction ~ 1 + Shots_on_five + Coach_Experience + Shots_on_five:Coach_Experience +
## (1 | Team_ID)
## Data: nba
##
##      AIC      BIC    logLik deviance df.resid
##    1348    1370     -668    1336      294
##
```

```
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -2.9627 -0.7635  0.0034  0.7916  2.6178
##
## Random effects:
##   Groups   Name      Variance Std.Dev.
##   Team_ID  (Intercept) 0.00     0.00
##   Residual                5.03     2.24
## Number of obs: 300, groups: Team_ID, 30
##
## Fixed effects:
##
##              Estimate Std. Error t value
## (Intercept)         4.491      0.726    6.19
## Shots_on_five        2.883      0.314    9.19
## Coach_Experience       1.247      0.412    3.03
## Shots_on_five:Coach_Experience  0.125      0.144    0.87
##
## Correlation of Fixed Effects:
##              (Intr) Shts__ Cch_Ex
## Shots_on_fv -0.825
## Coch_Exprnc -0.909  0.669
## Shts_n_:C_E  0.863 -0.921 -0.855
```

Using the fixed-effects estimates, the fitted equation for the fixed-effects model is:

$$\widehat{\text{Life Satisfaction}}_{ij} = 6.41 + 3.37(\text{SO5} + 1.55(\text{Coach Experience}))$$

Interpreting these coefficients,

- The predicted average life satisfaction for all players who successfully make 0 of their five shots, and whose coach has 0 years of experience is 4.49.
- The effect of shots-on-five on life satisfaction depends on the effect of coach experience. -The effect of coach experience on life satisfaction depends on the effect of shots-on-five.

In this model, there are two sources of variation (team intercepts, and random error), and thus two variance estimates. The variance estimates are:

- $\hat{\sigma}_{\epsilon}^2 = 5.03$
- $\hat{\sigma}_0^2 < .01$

Displaying the Results of the Fitted Models

It is common to display the results of the fitted models in a table or plot. Typically we would use the table to show the results of a series of fitted models and display the adopted “final” model(s) in a plot. For this example, I would display the results of the fitted models that included a random-effect of intercept in a table. (I probably would not display the models that included the random-effect of slope.) I would plot the results of the two top models from the table of evidence. (I would plot each model in a different panel.)

Table of Fitted Models

In displaying the results from fitted mixed-effects models, we typically provide (1) fixed-effects estimates; (2) variance component estimates; and (3) model-level evidence (e.g., LL, AIC). If you are using Markdown, there are several packages that can be used to obtain syntax for these types of tables. I use the **texreg** package for this, but you could also use **stargazer**.

```

library(texreg)

# Fit the models you want to present
lmer.0 = lmer(Life_Satisfaction ~ 1 + (1 | Team_ID), data = nba, REML = FALSE)

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

lmer.2 = lmer(Life_Satisfaction ~ 1 + Shots_on_five + Coach_Experience +
             (1 | Team_ID), data = nba, REML = FALSE)

lmer.3 = lmer(Life_Satisfaction ~ 1 + Shots_on_five + Coach_Experience +
             Shots_on_five:Coach_Experience + (1 | Team_ID), data = nba, REML = FALSE)

# Use extract.lmerMod() to remove number of observations
# for better table formatting
tr0 = extract.lmerMod(lmer.0, include.nobs = FALSE, include.groups = FALSE,
                     include.bic = FALSE)
tr1 = extract.lmerMod(lmer.1, include.nobs = FALSE, include.groups = FALSE,
                     include.bic = FALSE)
tr2 = extract.lmerMod(lmer.2, include.nobs = FALSE, include.groups = FALSE,
                     include.bic = FALSE)
tr3 = extract.lmerMod(lmer.3, include.nobs = FALSE, include.groups = FALSE,
                     include.bic = FALSE)

texreg(
  list(tr0, tr1, tr2, tr3),
  custom.coef.names = c("Intercept", "Player success", "Coach experience",
                        "Player success x Coach experience"),
  reorder.coef = c(2, 3, 4, 1),
  caption = "Coefficients (Standard Errors) for a Taxonomy of Fitted Mixed-Effects Models to Predict Life Satisfaction",
  caption.above = TRUE
)

```

The `texreg()` function outputs the syntax for a \LaTeX formatted table. (If you are outputting to HTML rather than PDF use `htmlreg()` instead of `texreg()` from the same package.) I often output the syntax for the \LaTeX formatted table, copy-and-paste this into the document (not in an R code chunk) and then change syntax to customize the table. For example, I often change the AIC values to AICc values; I also typically add rows to identify fixed-effects, random-effects, and model-level evidence; I often change the names of the random-effects to associate the level-2 structure more clearly (e.g., in this example, team); and lastly, I typically delete the stars indicating p -values (these p -values are contentious).

Table 5: Coefficients (Standard Errors) for a Taxonomy of Fitted Mixed-Effects Models to Predict Life Satisfaction for 300 NBA Players from 30 Teams. All Models are Fitted using Maximum Likelihood.

	Model 1	Model 2	Model 3	Model 4
<i>Fixed-effects</i>				
Player success		3.37 (0.12)	3.13 (0.12)	2.88 (0.31)
Coach experience			1.55 (0.21)	1.25 (0.41)
Player success x Coach experience				0.12 (0.14)
Intercept	14.81 (0.73)	6.41 (0.36)	3.95 (0.37)	4.49 (0.73)
<i>Variance estimates</i>				
Team (Intercept)	14.41	0.81	0.00	0.00
Residual	14.61	5.27	5.05	5.03
<i>Model-level evidence</i>				
Log Likelihood	-863.68	-688.92	-668.50	-668.13
AICc	1733	1386	1347	1349

Plot of “Final” Model(s)

If you plot the model results, it is typical to plot the fixed-effects part of the mixed-effects model (the average model; not by team). We do this exactly the same way we do for plotting the results from `lm()`. Here I use a sequence of values for shots-on-five (plotted on the x -axis) and a low and high value for coaching experience (based on the data)

```
# Set up plotting data for lmer.2
my_data2 = expand.grid(
  Shots_on_five = seq(from = 0, to = 5, by = 1),
  Coach_Experience = c(1, 3),
  model = "Main-effects"
)

# Set up plotting data for lmer.3
my_data3 = expand.grid(
  Shots_on_five = seq(from = 0, to = 5, by = 1),
  Coach_Experience = c(1, 3),
  model = "Interaction"
)
```

We then use the `predict()` function to obtain the \hat{Y} values for both models under consideration. We include the `re.form=NA` argument to ignore the random-effects (only compute the \hat{Y} values based on the fixed-effects).

```
# Predict life satisfaction
my_data2$yhat = predict(lmer.2, newdata = my_data2, re.form = NA)
my_data3$yhat = predict(lmer.3, newdata = my_data3, re.form = NA)

# Combine both data sets into one
my_data = rbind(my_data2, my_data3)
head(my_data)
```

Shots_on_five	Coach_Experience	model	yhat
0	1	Main-effects	5.50
1	1	Main-effects	8.63
2	1	Main-effects	11.77
3	1	Main-effects	14.90
4	1	Main-effects	18.04
5	1	Main-effects	21.17

```
# Turn coach experience into a factor for better plotting
my_data$Coach_Experience = factor(my_data$Coach_Experience)

ggplot(data = my_data, aes(x = Shots_on_five, y = yhat, color = Coach_Experience)) +
  geom_line() +
  theme_bw() +
  xlab("Shots-on-five") +
  ylab("Predicted life satisfaction") +
  ggsci::scale_color_d3(name = "Coach experience (in years)") +
  facet_wrap(~model)
```

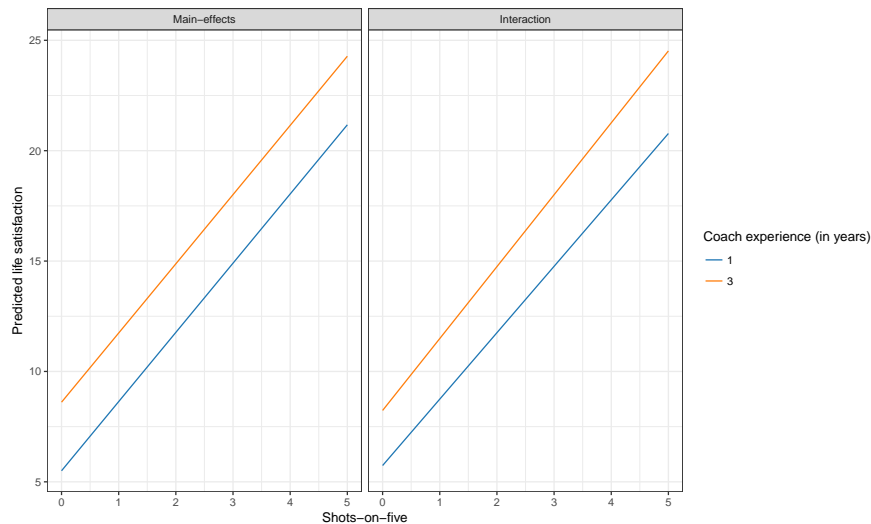


Figure 3: Plot of the predicted life satisfaction scores as a function of shots-on-five based on the main-effects with coach experience (left-side) and interaction with coach experience (right-side) models. Predicted values are shown for one year of coaching experience (blue lines) and three years of coaching experience (orange lines)

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.