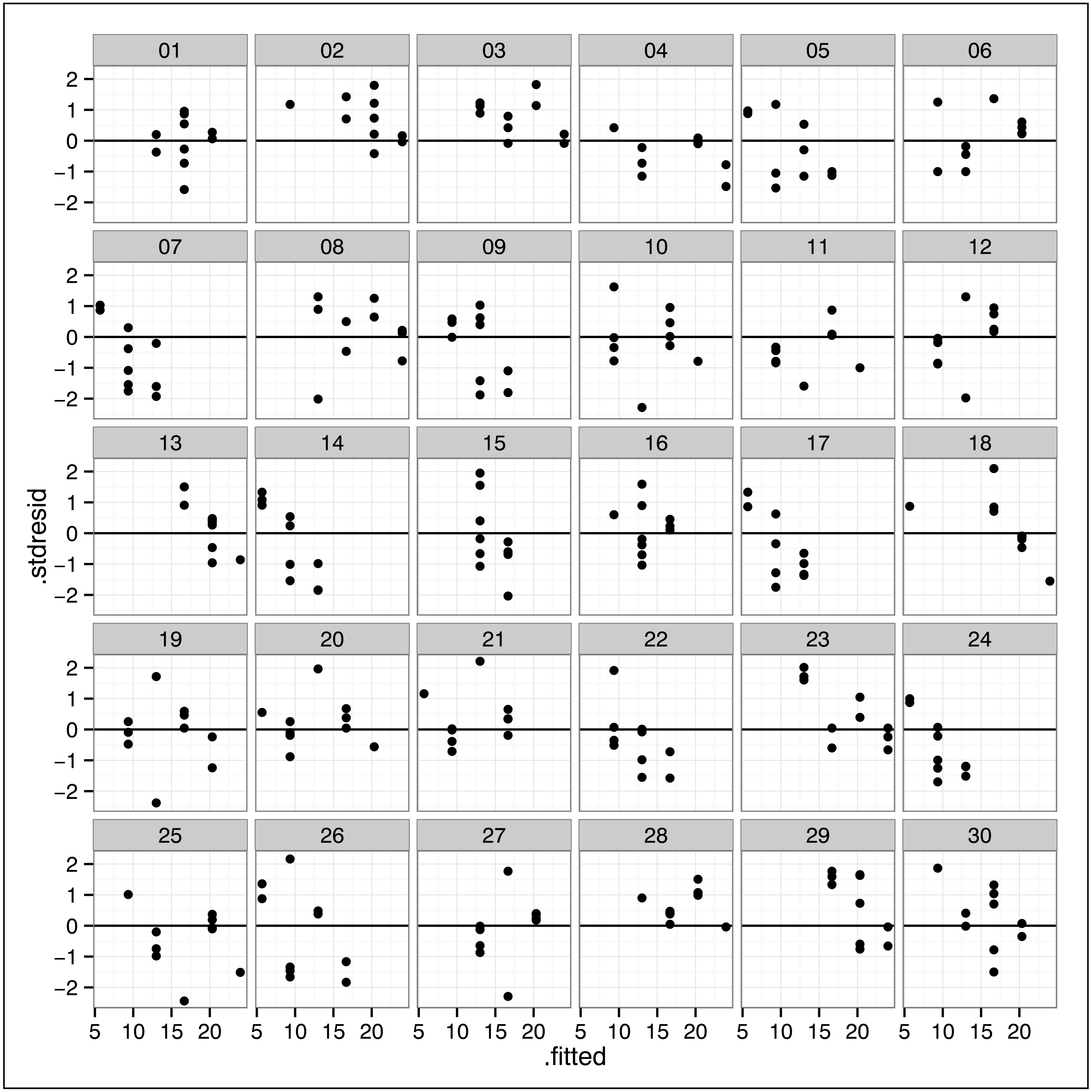
Introduction to Mixed-Effects Models

Fit LM at the Case-Level

```
Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)     5.6937     0.3026   18.81   <2e-16 ***
Shots_on_five   3.6598     0.1075   34.04   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2.444 on 298 degrees of freedom
Multiple R-squared:  0.7954,   Adjusted R-squared:  0.7948
F-statistic:  1159 on 1 and 298 DF,  p-value: < 2.2e-16
```

Differences in free-throw shooting (Shots_on_five) seem to explain differences in life satisfaction, $F(1, 298) = 1159$, $p < .001$. These differences explain 79.5% of the variation in life satisfaction scores.



Within teams, the residuals are either *mostly positive* or *mostly negative*

This is a violation of the independence assumption.

Why is this a Problem?

```
Analysis of Variance Table

Response: Life_Satisfaction
               Df Sum Sq Mean Sq F value    Pr(>F)
Shots_on_five   1 6924.5  6924.5  1158.8 < 2.2e-16 ***
Residuals     298 1780.7     6.0
---
```

300 independent observations = 299 *df*

Model 1 *df*

Residual 298 *df*

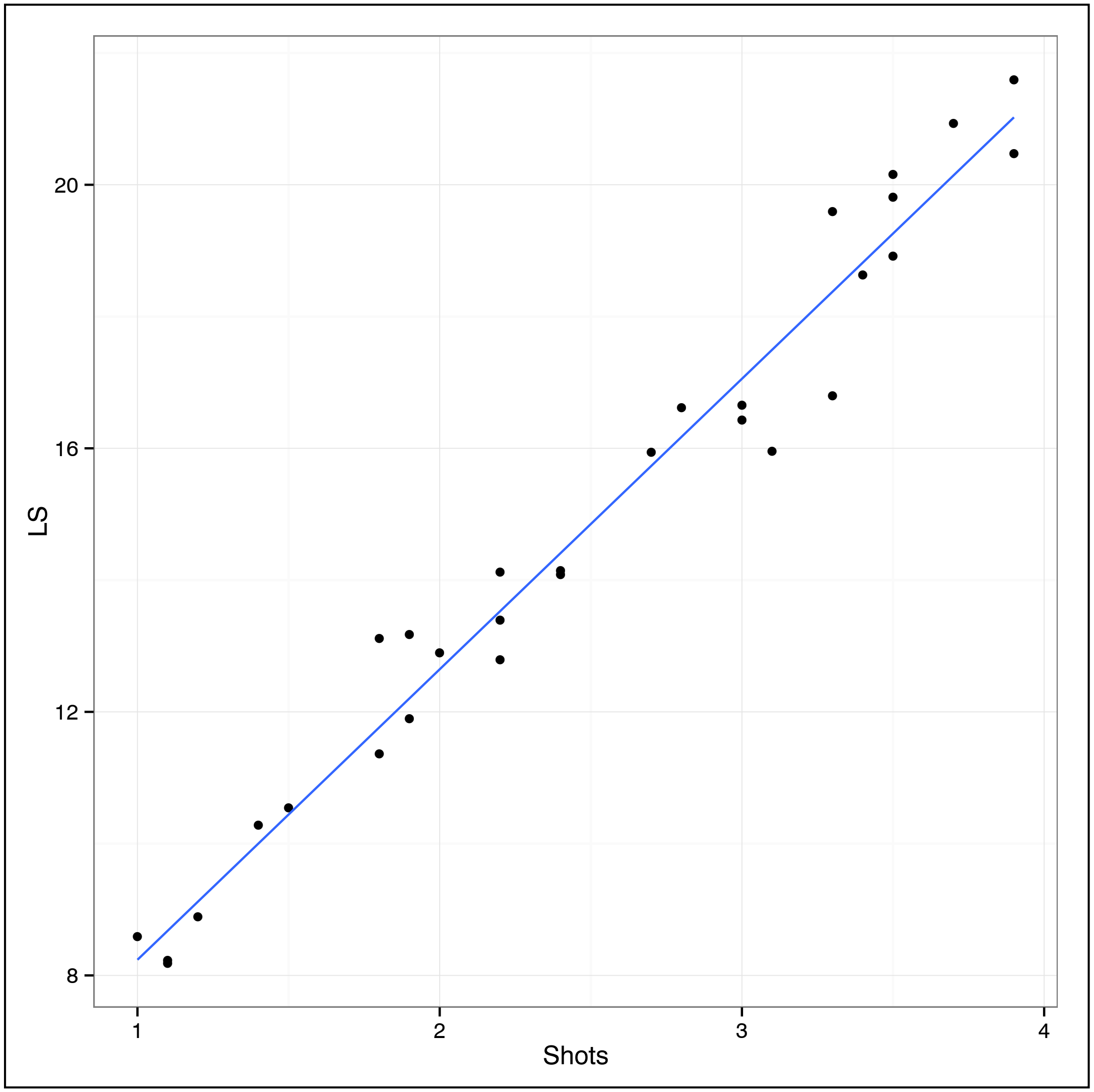
$$F = \frac{\mathrm{MS}_{\mathrm{Model}}}{\mathrm{MS}_{\mathrm{Residual}}}$$

$$F = \frac{\left(\dfrac{\mathrm{SS}_{\mathrm{Model}}}{df_{\mathrm{Model}}}\right)}{\left(\dfrac{\mathrm{SS}_{\mathrm{Residual}}}{\boxed{df_{\mathrm{Residual}}}}\right)}$$

Violation of the independence assumption indicates that it is wrong to think that the data constitutes 300 independent observations….which means that we have fewer *df* for the residual than we think...

Solution 1: Analysis on the Group Means

Rather than fitting the model to the 300 cases, we can fit it to the 30 team-level observations.



Fit LM at the Group-Level

```
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   3.8283     0.3912   9.787 1.55e-10 ***
Shots         4.4090     0.1479  29.802  < 2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.7209 on 28 degrees of freedom
Multiple R-squared:  0.9694,	Adjusted R-squared:  0.9683
F-statistic: 888.2 on 1 and 28 DF,  p-value: < 2.2e-16
```

Differences in team-level free-throws seem to explain about 97% of the variation in differences in the team mean life satisfaction scores, $F(1, 28) = 888.2$, $p < .001$.

Note that R^2 is higher in this analysis since we it only considers the *between-team variation* and **not** the *within-team variation*.

The *residual df*, which now reflects the more reasonable assumption that teams cannot be treated independently, is much lower than in the previous analysis (28 vs. 198). Consequently, the p-value is higher, albeit still statistically reliable.

The estimated coefficients are not that different from the initial analysis.

Two Problems with the Group-Level Analysis

- We do not account for the variation in life satisfaction scores within a particular team
- We do not account for variation in the number of players sampled within each team. In the group-level analysis, each team is weighted equally, even though the averages in practice could be based on more players on some teams than others (in this example $n = 10$ for each team).

Solution: Fit a Multi-Level Model

Level-1: $\text{life satisfaction}_{ij} = \beta_0^* + \beta_1^*(\text{Shots on five}_{ij}) + \epsilon_{ij}$

Level-2: $\beta_0^* = \beta_{00} + \eta_{0j}$

$\beta_1^* = \beta_{01} + \eta_{1j}$

The **level-1 model** specifies how the player-level predictors relate to the outcome (life satisfaction scores)

The **level-2 model** specifies how each coefficient in the level-1 model is predicted by team-level predictors

This is called a multi-level model.

Level-1 Model

$$\text{life satisfaction}_{ij} = \beta_0^* + \beta_1^*(\text{Shots on five}_{ij}) + \epsilon_{ij}$$

- Level-1 model specifies the regression equation across players
  - ✓ β_0 is the intercept across players
  - ✓ β_1 is the slope across players
  - ✓ ε_{ij} are the random errors for the *i*th player to the regression line

- Only source of variation in Level-1 model is player-to-player (within-team) variation
- Any unaccounted for player-to-player variation is absorbed into the residual, ε_{ij}

Level-2 Model

$$\beta_0^* = \beta_{00} + \eta_{0j}$$

$$\beta_1^* = \beta_{01} + \eta_{1j}$$

- Level-2 models are team-level regression models
  - ✓ Outcome variables are the Level-1 regression parameters (e.g., intercepts and slopes)
  - ✓ Number of Level-2 equations determined by number of regression parameters in the Level-1 model
  - ✓ Covariates that account for team-to-team (between-team) variation exclusively appear in Level-2 model

- Error terms in Level-2 model are called **random-effects**
  - ✓ Can only be as many random-effects as there are Level-2 equations
  - ✓ Not every Level-2 equation needs to have random-effect

Mixed-Effects Model

Level-1: $$\text{life satisfaction}_{ij} = \beta_0^* + \beta_1^*(\text{Shots on five}_{ij}) + \epsilon_{ij}$$

Level-2: $$\beta_0^* = \beta_{00} + \eta_{0j}$$

$$\beta_1^* = \beta_{01} + \eta_{1j}$$

Substitute the Level-2 equations into the Level-1 equation to get the mixed-effects model

$$\text{life satisfaction}_{ij} = [\beta_{00} + \eta_{0j}] + [\beta_{01} + \eta_{1j}]\,(\text{Shots on five}_{ij}) + \epsilon_{ij}$$

The model essentially helps us to partition the variation in the residual...that which is due to between-school (school-to-school) variation (the random effects) and that which is within-school (student-to-student) error.

$$\text{life satisfaction}_{ij} = \beta_0^* + \beta_1^*(\text{Shots on five}_{ij}) + \boxed{\epsilon_{ij}}$$

$$\text{life satisfaction}_{ij} = \beta_{00} + \beta_{01}(\text{Shots on five}_{ij}) + \boxed{\eta_{0j} + \eta_{1j}(\text{Shots on five}_{ij}) + \epsilon_{ij}}$$

The mixed-effects model is used to specify the R syntax for fitting the model

$$\text{life satisfaction}_{ij} = \beta_{00}(1) + \beta_{01}(\text{Shots on five}_{ij}) + \eta_{0j}(1) + \eta_{1j}(\text{Shots on five}_{ij}) + \epsilon_{ij}$$

```
> library(lme4)
> lmer.1 = lmer(Life_Satisfaction ~ 1 + Shots_on_five +
    (1 + Shots_on_five | Team_ID), data = nba)
```

Level-2 ID

$$\text{life satisfaction}_{ij} = \beta_{00}(1) + \beta_{01}(\text{Shots on five}_{ij}) + \eta_{0j}(1) + \eta_{1j}(\text{Shots on five}_{ij}) + \epsilon_{ij}$$

The `lmer()` function uses syntax similar to the `lm()` function.

- The fixed-effects use the exact same syntax
- The random-effects appear in parentheses and reference the level-2 ID variable

```
> summary(lmer.1)

Fixed effects:
              Estimate Std. Error t value
(Intercept)     6.4296     0.3169   20.29
Shots_on_five   3.2887     0.1340   24.55
```

The fixed-effects are interpreted exactly the same as we interpret the coefficients in an LM.

- The predicted life satisfaction score for a player who shoots zero free-throws is 6.42.
- The average difference in life satisfaction scores between students who are one free-throw different is predicted to be 3.29.

There are no *p*-values given for results from a mixed-effects analysis in R. That is because it is theoretically unclear what value for the *df* should be used for the *t*-test.

The output associated with the random-effects provide variance estimates for the associated random-effects in the model and also for the residuals.

```
Random effects:
 Groups   Name          Variance Std.Dev. Corr
 Team_ID  (Intercept)   0.09279  0.3046
          Shots_on_five 0.09913  0.3148   1.00
 Residual               5.10616  2.2597
Number of obs: 300, groups:  Team_ID, 30
```

Level-1 variance: Quantifies the amount of life satisfaction variation **within** teams

Level-2 variance: Quantifies the amount of variation in life satisfaction **between** teams.

$$\epsilon_{ij} \sim \mathcal{N}\left(0, \hat{\sigma}^2_{\epsilon}\right) \qquad \hat{\sigma}^2_{\epsilon} = 5.106$$

$$\beta_{0j} \sim \mathcal{N}\left(\beta^*_0, \hat{\sigma}^2_{\beta_{0j}}\right) \qquad \hat{\sigma}^2_{\beta_{0j}} = 0.093$$

$$\beta_{1j} \sim \mathcal{N}\left(\beta^*_1, \hat{\sigma}^2_{\beta_{1j}}\right) \qquad \hat{\sigma}^2_{\beta_{1j}} = 0.099$$

The level-2 (between-team) variance estimates tell us how much the school-level intercept and slopes vary.

By partitioning the error variance, we get a more appropriate **residual variance** (MS) on which to test the effect of SES.

```
> anova(lmer.1)

Analysis of Variance Table
              Df Sum Sq Mean Sq F value
Shots_on_five  1 3077.6  3077.6  602.72
```

$$\hat{\sigma}_\epsilon^2 = 5.106 \qquad F = \frac{3077.6}{5.106} = 602.7$$

Again, no *p*-values given for the same reason...we have no idea what the appropriate *df* are.

We could test this if we knew which *df* to use in the denominator….298? 28? somewhere in between?

```
# Best case scenario (assume independence)
> 1 - pf(602.7, df1 = 1, df2 = 298)

[1] 0


# Worst case scenario
> 1 - pf(602.7, df1 = 1, df2 = 28)

[1] 0
```

In this example it doesn't matter, but in practice when the sample size and the number of groups is smaller this can make a world of difference.

Testing Fixed Effects in Mixed-Effects Models

$$t = \frac{\hat{\beta}_j}{\mathrm{SE}_{\hat{\beta}_j}}$$

Fixed effects in multilevel regression are typically tested by creating a ratio of the slope/intercept estimate to the estimate of the standard error. This kind of ratio, the Wald ratio, usually distributed as a *z* or *t*, is used in many statistical tests.

In the HLM software, the *df* associated with the *t*-distribution is computed as

$$df = N - q - 1$$

where N is the number of level-2 units (groups), q is the number of predictors in the model.

In practice, this formula does not seem to be used precisely, and if you re-run the analyses several times, you will notice somewhat different degrees of freedom listed in the output under—**approximate df**. In addition, test of the effects of cross-level interactions often use degrees of freedom based on the number of level-1 units (i.e., total number of individuals in the sample).

Testing Fixed Effects in Mixed-Effects Models

In SPSS, the same ratio is called z—a Wald test. SPSS uses a Satterthwaite adjustment to the *df* based on the number of level-1 cases. Satterthwaite degrees of freedom are a way of proportionally adjusting the *df* to provide a more accurate *p*-value estimate from the family of distributions.

SAS has many options for *df* adjustment. The Kenward–Rogers adjustment makes further adjustments based on the fact the the values used in the variance–covariance matrix are estimates and not known quantities.

Note. For samples in which the number of cases at both level-1 and level-2 are large, all methods give almost equivalent results. For small samples, it is recommended that the Kenward–Rogers method be used (Verbeke & Molenberghs, 2000).

The **lmerTest** package can be used to get the Kenward-Rogers estimated *p*-values. In this class, I don't want to see any *p*-values. We will use information criteria to evaluate models. To evaluate predictors, it is just as good to use the heuristic of whether $t > 2$.

Obtaining *p*-values for Fixed-Effects in R

```
# From the model summary() output
REML criterion at convergence: 1379

> AIC(lmer.1)
[1] 1391.018

> library(MuMIn)
> AICc(lmer.1)
[1] 1391.305
```

The default estimation method for `lmer()` is REML. The AIC and AICc values are based on the REML values for the log-likelihood. To fit a model using ML we include `REML=FALSE` to the function as an additional argument.