

1 Pooled/unpooled v.s. fixed/random effects

You may have noticed that we use a couple of different terms interchangeably in this class when it comes to models. Sometimes we talk about coefficients as being completely pooled/partially pooled/unpooled, and sometimes we talk about coefficients as being random or fixed. Yikes, so confusing! Here's a quick document explaining what these various terms mean and what sorts of models they represent. We're only going to be talking about models where the pooling applied to the intercept and slope is the same; most models look like this, and these models are easier to talk about. You should be able to see how you might pool different coefficients differently, though the R code for that can be challenging. We'll use the HSB data, and all of the models we'll consider will look at regressions of math achievement on SES.

1.1 Completely pooled

A completely pooled model is a model where we assume that every second-level unit (school) has the same intercept and slope (slopes and intercepts are both completely pooled). This doesn't really have an analog in the fixed/random effects world.

A completely pooled model in this setting might look like

$$\begin{aligned} \text{mathach}_i &= \beta_0 + \beta_1 \text{SES}_i + \varepsilon_i \\ \varepsilon_i &\sim \text{Normal}(0, \sigma^2) \end{aligned}$$

In a completely pooled model we're basically assuming that every school has the same intercept and slope, so we just ignore school membership; notice that we don't even include the j subscript because we're ignoring schools completely. How rude!

We would fit this model with

```
lm(mathach ~ ses).
```

1.2 Partially pooled

A partially pooled model allows for the possibility that different schools might have different slopes and intercepts, but assumes that these slopes and intercepts come from a Normal distribution, which has the effect of pulling them all in towards a grand mean (or *partially pooling* them). This model can also be called a model with random slopes and random intercepts, since we assume that school intercepts and residuals are random draws from a multivariate distribution with means equal to the grand means (and some possibly non-0 correlation). We don't try to estimate these by themselves, only their variances and covariance.

This model can be represented as

$$\begin{aligned} \text{mathach}_{ij} &= \beta_{0j[i]} + \beta_{1j[i]} \text{SES}_i + \varepsilon_i, \\ \beta_{0j} &= \gamma_{00} + u_{0j}, \\ \beta_{1j} &= \gamma_{10} + u_{1j}, \\ \varepsilon_i &\sim \text{Normal}(0, \sigma_\varepsilon^2) \\ \begin{pmatrix} u_{0j} \\ u_{1j} \end{pmatrix} &\sim N \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_0^2 & \rho\sigma_0\sigma_1 \\ \rho\sigma_0\sigma_1 & \sigma_1^2 \end{pmatrix} \right] \end{aligned}$$

We would fit this model with

```
lmer(mathach ~ ses + (ses|school)).
```

1.3 Unpooled

In an unpooled model, we don't share *any* information across schools about the slopes and intercepts. Instead, we estimate each one separately in each higher-order unit. This is a *fixed-effects* model, because the model treats each school-level slope and intercept as a fixed quantity in the population to be estimated directly. In general parlance, to be a little more precise, a fixed-effects model is a model with unpooled intercepts and completely pooled slopes (although in theory the completely pooled model also has only fixed effects, it's just that those effects are the same in every school; this is why the language of completely pooled, partially pooled, and unpooled coefficients is a little more precise, though it's also less popular).

We could represent an unpooled model as

$$\begin{aligned} \text{mathach}_{ij} &= \beta_{0j[i]} + \beta_{1j[i]}SES_i + \varepsilon_i \\ \varepsilon_i &\sim \text{Normal}(0, \sigma^2) \end{aligned}$$

We would fit the model with

```
lm(mathach ~ ses * school),
```

although we might get our estimates in a more useful way by specifying an (identical) model which has no reference school, i.e.,

```
lm(mathach ~ -1 + school + ses : school).
```

For either of these models to fit you need to ensure that school is coded as a factor and not a number; this is *not* a concern for `lmer`.