증거기반연구 6차 세미나

Week 3: HLM (위계적 선형 모형)

주상원 (sangwonju.netlify.app)

today

열심히 해봅시다… 영차영차…

목차

i. **이론**

- 위계적 선형모형 (Hierarchical Linear Modelling = Mixed Effect Model = Multilevel Model)의 개념

> 핵심어: Gauss-Markov Theorem, Nested Data, Ecological fallacy

- HLM의 유형 및 분석단계별 유의사항

> 핵심어: Intercept and Slopes, MLE, ICC, AIC(BIC), Deviance, Fixed and Random Effects, Within vs. Between

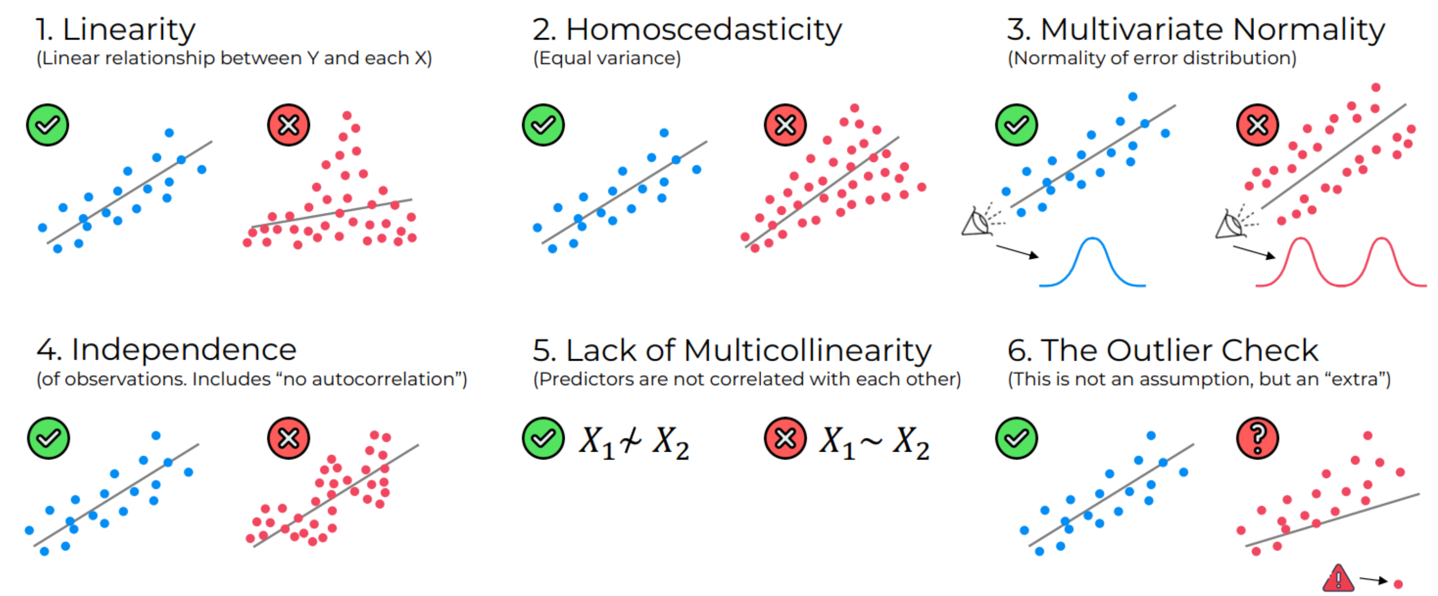
- Centering, Inter-Level Interaction

ii. **실습:** R과 Stata로 실제 데이터 분석

Module I: 위계적 선형모형의 개념

회귀분석의 기본 가정s

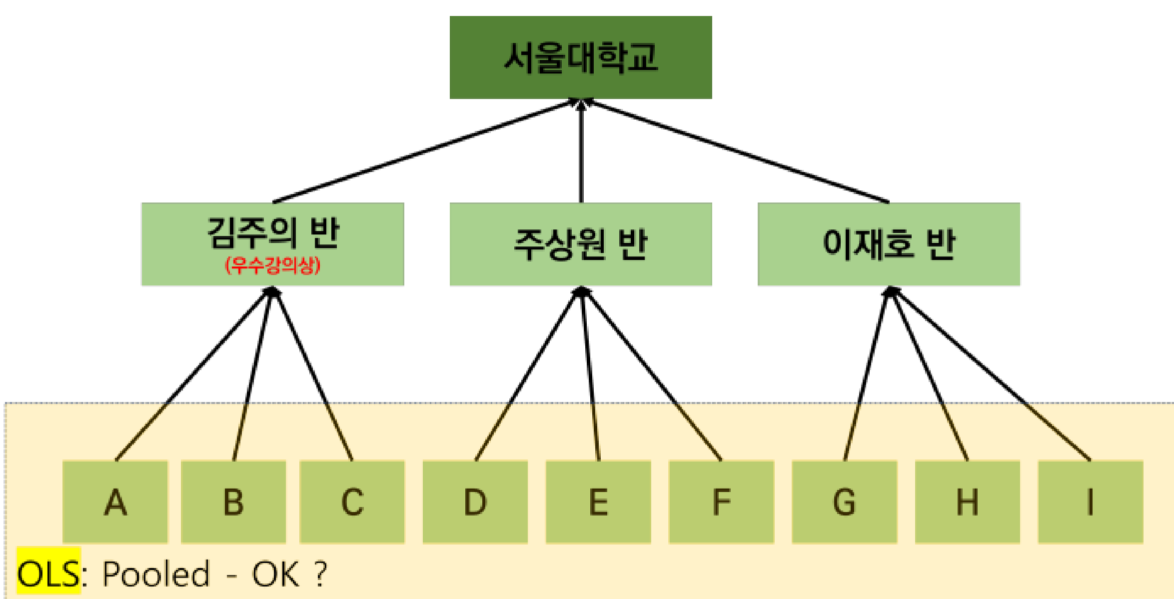
- **Gauss-Markov Theorem**: 선형회귀분석에서 (1) 선형이고, (2, 3) 오차(ϵ)가 $\epsilon \sim N(0, \sigma^2)$, (4) 오차가 상관관계가 없고, (5) 설명변수가 외생변수일 때 최소제곱 추정량(OLS)은 BLUE(Best **Linear Unbiased** Estimator)이다. 출처 (https://www.econometrics-with-r.org/5-5-the-gauss-markov-theorem.html)



Assumptions

그러나 현실은...? 특히 조직연구에서는..?

- Nested Data: 학교 > 학년 > 반 > 개인
- 조직연구의 경우: e.g. 공무원 조직: 공무원 개인 → 과 → 국 → 처



위계적 다층자료 (Cont’d)

- 위계구조(hierarchical structure)
  - 하나의 단위가 그보다 상위 수준(level)에 속해 있는 구조
  - 데이터의 위계는 연구자의 목적에 따라 여러 단계로 계층화 할 수도 있고, 여러 수준으로 구 분하여 구축할 수 있음

위계적 다층자료: Nested Data의 주요 형태

- 위계적 구조: Long Form vs. Wide Form

Wide Form Nested {#tbl-first}

이름	1반	2반	3반	통계성적
A	0	0	0	95
B	1	0	1	65
C	0	1	0	85
D	0	1	0	45
E	0	0	1	75

Long Form Nested {#tbl-second}

이름	소속반	통계성적
A	1반	95
B	3반	65
C	2반	85
D	2반	45
E	3반	75

학생이 학교에 내재한 2-수준 다층자료의 예시

- *if 소속반을 시간으로 바꾸어 반복측정?:* **종단(Longitudinal) Panel Data**

위계적 다층자료 (Cont'd)

- **학생들은 각 학교에 내재한 구조**
  - 동일 반에 속한 학생들은 학교의 문화, 학습환경, 친구관계, 교사 등 수많은 요인들을 공유하므로 동일학교 학생들의 행동 (e.g. 성취도, 학습동기)에 영향을 미침
- **상호의존성**
  - 개인을 둘러싼 지역의 경제, 사회, 문화, 제도 및 물리적 특성이 개개인의 행동, 선호 도, 가치관 등에 영향을 미치기 때문
  - 관찰단위인 학생들은 동일 학교, 동일 반 내에서는 **상호의존성**을 가지게 되고, 소속이 다르면 독립성을 갖음 → 독립성 가정이 결과적으로 위배되게 됨
- **복수의 Unit of Analysis**
  - 학생수준에서의 변수 (Level 1), 학교수준에서의 변수 (Level 2)

- **불균형 자료**
    - 각 학교별로 학생이 다르기에 관측치는 일반적으로 같을 수 없음

독립성 (Independence) 가정의 위배

i. **Error와 독립변수간에 상관관계가 없어야 함 (Endogeneity issue)**

- 중요한 변수가 모형에서 생략되어지거나, 비체계적 오류로 인한 측정오차, 독립변수와 종속변수간 동시상관시 발생

ii. **Error 간에 상관관계가 없어야 함**

- $COV(e_i, e_j) = E(e_i e_j) - E(e_i)E(e_j) = 0$
  상위수준(Level 2) 군집화 $\rightarrow$ Error간에 상관관계가 발생한다면?
  회귀분석의 추정치의 분산이 과도하게 커짐 $\rightarrow$ 정확 X

독립성 (Independence) 가정의 위배 (Cont'd)

- 사례 분석: **학습시간과 통계과목 성적간의 관계 분석**
    - 데이터

기술통계량

skim_type	skim_variable	n_missing	complete_rate	factor.ordered	factor.n_unique	factor.top_counts	numeric.mean	numeric.sd	numeric.p0	numeric.p25	numeric.p50	numeric.p75	numeric.p100	numeric.hist
factor	group	0	1	FALSE	30	1: 100, 2: 100, 3: 100, 4: 100	NA	NA	NA	NA	NA	NA	NA	NA
numeric	x	0	1	NA	NA	NA	40.62740	6.759487	22.89716	35.13338	40.44313	45.90419	64.73969	▁▃▇▇▂▁
numeric	y	0	1	NA	NA	NA	66.30165	14.964766	36.70912	53.70978	66.29439	79.20288	97.40393	▅▇▇▇▇▃

Raw Data

```
simul %>%
  transmute(group = factor(group),
            x,
            y) %>%
  mutate(across(where(is.numeric), ~round(.,3))) %>%
  DT::datatable(., options = list(pageLength = 5))
```

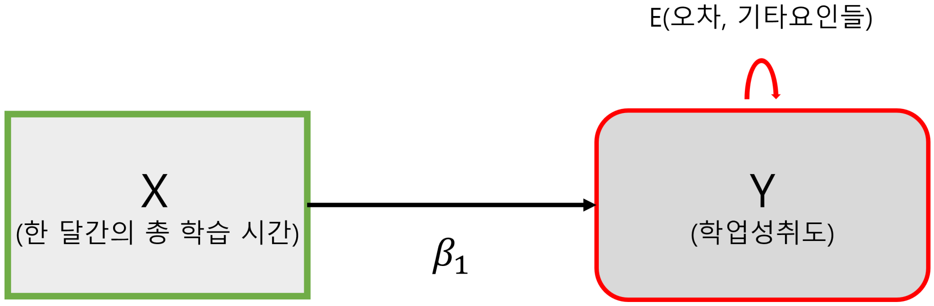
Show 5 entries Search:

	group	x	y
1	1	43.7	40.661
2	1	48.834	43.368
3	1	47.474	42.623
4	1	52.716	43.324
5	1	46.622	38.868

Showing 1 to 5 of 3,000 entries

Previous 1 2 3 4 5 … 600 Next

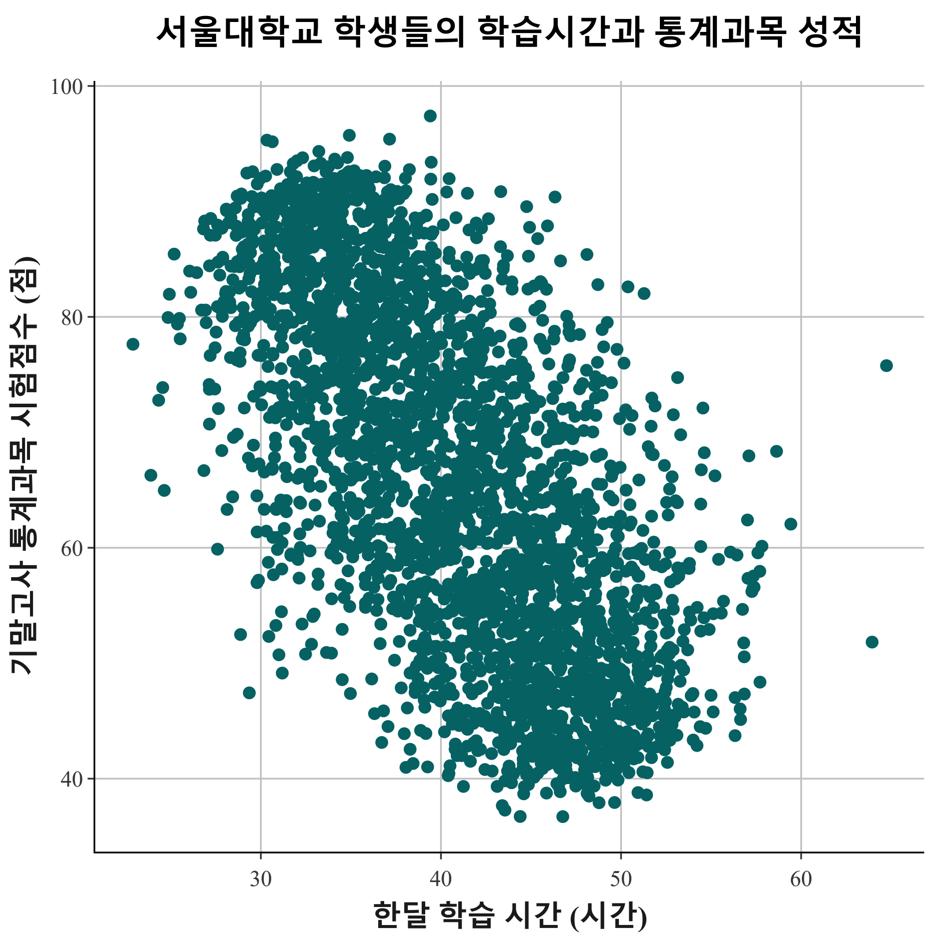
- 모형



$$Y = \beta_0 + \beta_1 X + \epsilon_i, \qquad \epsilon \sim N(0, \sigma^2)$$

독립성 (Independence) 가정의 위배 (Cont'd)

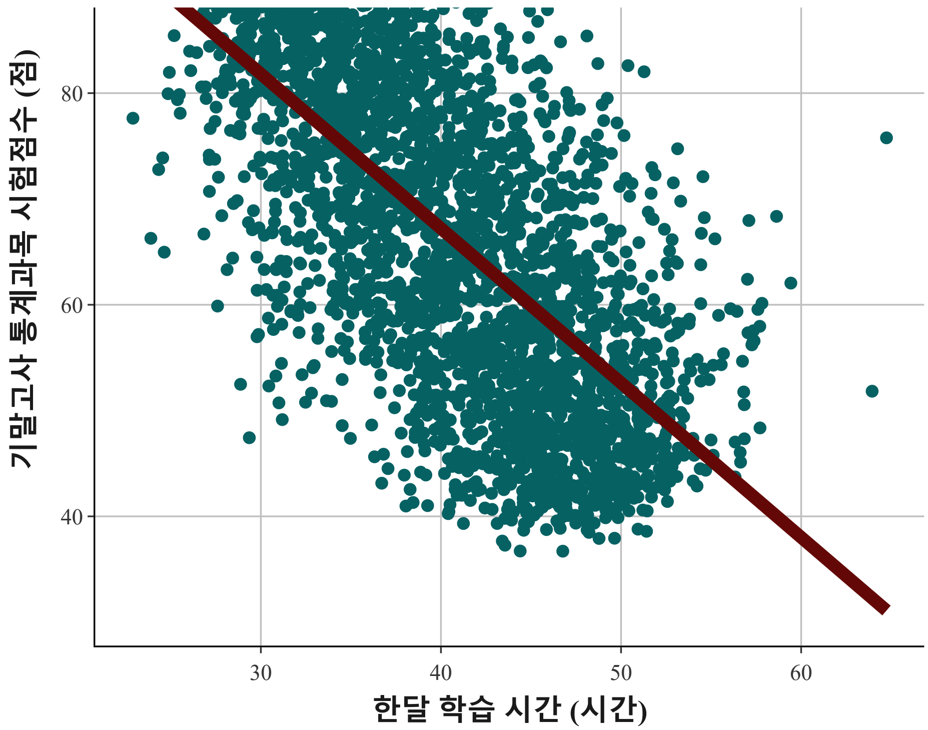
- 시각화



독립성 (Independence) 가정의 위배 (Cont'd)

- 시각화

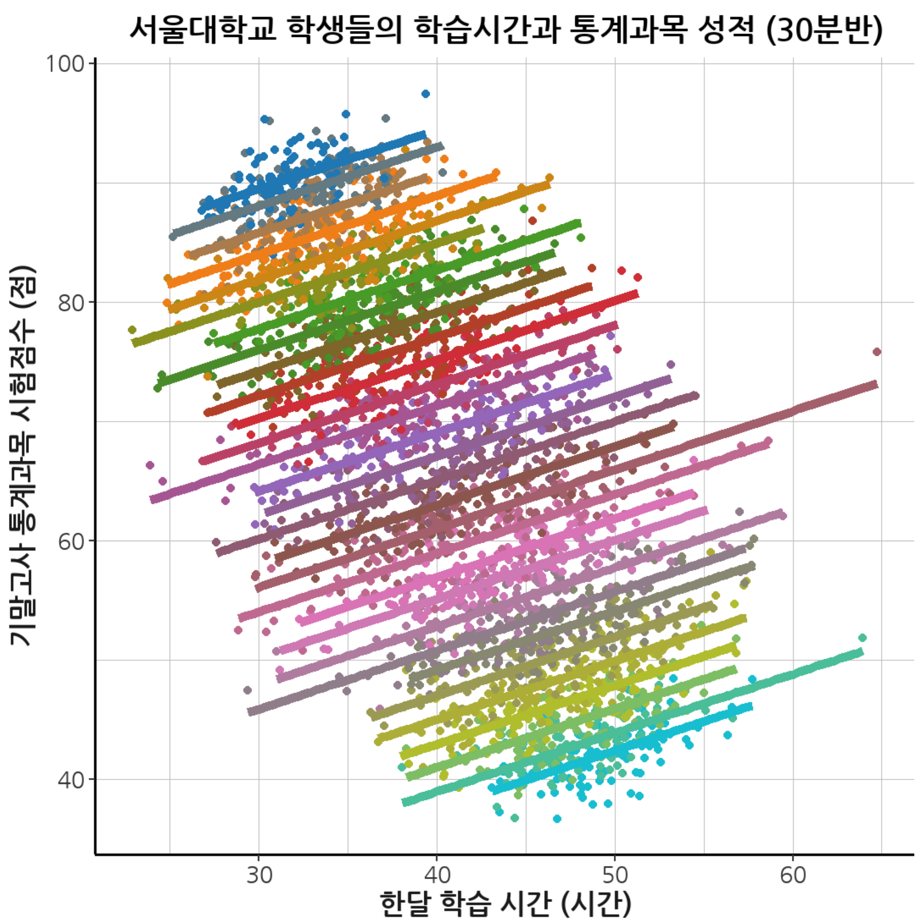
```
`geom_smooth()` using formula = 'y ~ x'
```



독립성 (Independence) 가정의 위배 (Cont'd)

- 시각화

```
`geom_smooth()` using formula = 'y ~ x'
```



독립성 (Independence) 가정의 위배 (Cont'd)

- **Pooled OLS**

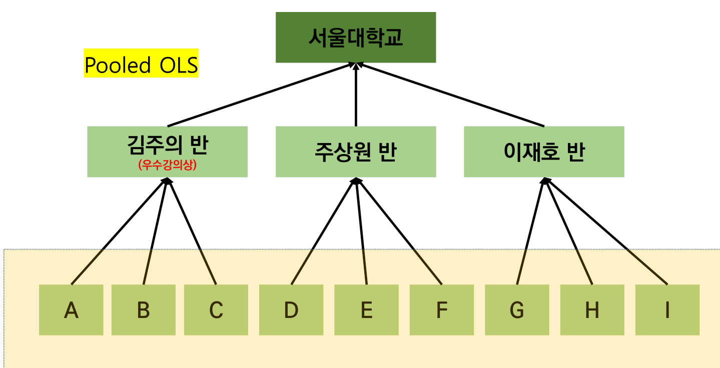
term	estimate	std.error	statistic	p.value
(Intercept)	125.613	1.252	100.337	0
x	-1.460	0.030	-48.028	0

- **Grouped regression** 30개 그룹의 estimates들의 평균

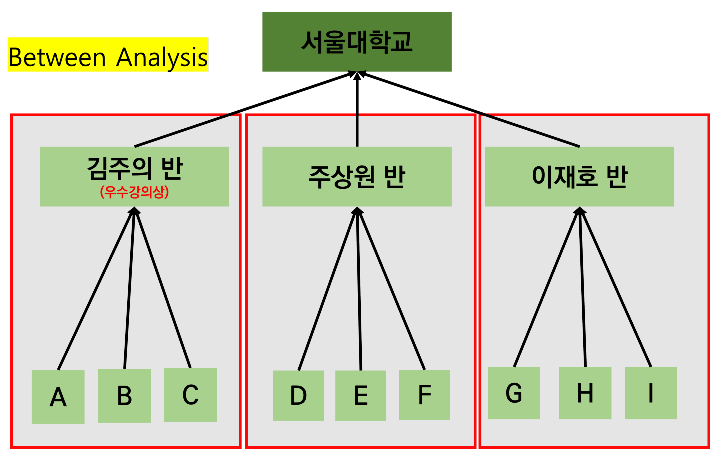
term	estimate	std.error	statistic
(Intercept)	46.263	1.941	25.361
x	0.493	0.048	10.775

기존 접근의 한계

- How to investigate relationships between variables that reside at different hierarchical levels (Bryk & Raudenbush, 2002)
- **Disaggregate?**: POLS → independence of obs assumption violated
- **Aggregate?**: Between → waste information from indvs, sample ↓, 생태학적 오류(Ecological Fallacy)



POLS



Between

Module I: Sum-up

- 현대 사회에서 인간의 다양한 위계에 nested 되어있는 존재이다.
- 이러한 조직 내에서의 군집성으로 인해 서로서로에게 영향을 주게 되고, 결과적으로 개체들 간의 독립성을 가정하는 가우스-마르코프 가정을 충족하는 것이 까다로움
  - 개개인간의 독립성이 존재하기 어렵고, 자료 자체가 nested되어 있다는 것을 고려하지 못함 → Individual 대상 연구에서 (특히, 조직맥락) OLS는 더이상 만능 X, 생태학적 오류를 범할 수 있게 됨
- Pooled-OLS는 집단간 군집성 무시, 가우스 마르코브 가정 위배, 심슨의 패러독스, Between 모형은 개개인들의 특성이 고려되지 않는 문제, 생태학적 오류
- **HLM의 필요성**: within과 between을 동시에 다루고, level 1과 2를 동시에 고려

Module II: HLM의 개념

위계적 선형 모형 (HLM)

대안: Hierarchical Linear Models

- designed to overcome the weakness of the disaggregated and aggregated approaches

1. explicitly model both **individual and group level residuals**, therefore, recognizing the partial interdependence of individuals within the same group (compared to OLS)
2. investigate both **lower level unit and higher level unit** variance in the outcome measure

∴ Model both within and between group variance (i.e., able to preserve potentially meaningful within group variance) + Investigate the influence of higher level units on lower level outcomes

HLM의 장점

- **Improves estimation of individual effects**
- **Models cross-level effects**: an interaction
- **Better partitioning of variance and covariance** you have variance and covariance of data set, and you can think about how much is due to Level 2 and how much is due to Level 1 etc. E.g. How much is school, how much is student?
- **No assumption of homogeneity of slopes** i.e., that each data entry can have a different slope
- **No assumption of independence** because in this model they are correlated
- **Missing data OK** the structure of the data – you don't need data for people at every time point (e.g., repeated measures), or every group has to have a score for every person (e.g., nested).

HLM의 이론

- **OLS**

$$Y_i = \beta_0 + \beta_1 X_i + \epsilon_i \quad \epsilon_i \sim N(0, \sigma^2)$$

- **HLM**은 집단내(개인: i)과 집단간(집단: j) 모형을 별개로 model specification 한다
  - **Level 1:**

$$Y_{0j} = \beta_{0j} + \beta_{1j} X_{ij} + \epsilon_{ij} \quad \epsilon_i \sim N(0, \sigma^2)$$

  - **Level 2:**

$$\beta_{0j} = \gamma_{00} + u_{0j}, \quad u_{0j} \sim N(0, \tau_{00})$$

$$\beta_{1j} = \gamma_{10} + u_{1j}, \quad u_{1j} \sim N(0, \tau_{11})$$

- **Fixed effects:**
  γ_{00} = average outcome for sample of groups
  γ_{10} = average individual effect (slope) on outcome
- **Random effects:**
  u_{0j} = unique effect of group j on average outcome
  u_{1j} = unique effect of group j on average slope

HLM을 통해 해결가능한 연구문제 examples

1. 우리나라 중학생 수학성취도의 학교간 교육격차는 어느정도인가?
2. 중학교 학생들의 수학성취도는 동일 학교 내에서 어느정도 차이가 있는가?
3. 학생들 사이의 수학성취도 차이에서 몇 %가 소속학교의 영향인가?
4. 학생들의 수학성취도는 개인차 요인의 영향을 더 많이 받는가 아니면 학교차 요인의 영향을 더 많이 받는가?
5. 저소득층 학생비율이 높은 학교의 학생들은 수학성취도에서 어느정도의 불이익을 받는가?
6. 학생 가정의 SES는 수학성취도와 어느정도 관련이 있는가?
7. 학생가정의 SES를 통제한 이후엥도, 학교별 고정평균 수학성취도는 여전히 학교간에 차이가 있는가?

8. 가정의 SES가 수학성취도에 미치는 효과는 모든 학교에서 유사한가? 만일 학교에 따라 다르다면 그 크기는 어느 정도인가?
9. 학생의 가정배경을 통제한 이후에 어떠한 특성의 중학교에서 평균 수학성취도가 높은가?
10. 학생들의 수학성취도가 가정환경에 영향을 받는 정도는 어떤 특성의 학교에서 더 커지는가?

Fixed Effect and Random Effect

In Panel, only on intercepts

- Fixed Effect: Time invariant Observation's Effect vs
- Random Effect: Time invariant Observation's Effect + *Random Part*

In HLM, similar but also on slopes

- **Fixed Effects**: parameter estimates that **do not vary across groups** (group invariant)
  The γ's in equations represent fixed effects
- **Random coefficients**: parameter estimates that are allowed to vary across groups such as the level-1 regression coefficients (e.g., β_{0j} and β_{1j}).

Fixed Effect and Random Effect (cont'd)

- **Fixed Effects**: 선형 회귀모델과 같이 절편 및 기울기의 추정계수가 집단에 따라 변화하지 않고 단 하나의 값을 가짐
- **Random Effect**: 절편 및 기울기의 추정계수가 하나의 값이 아니라 상위 수준인 집단의 특성에 따라 여러 개의 값을 가짐 (e.g., β_{0j} and β_{1j}).
- Equation

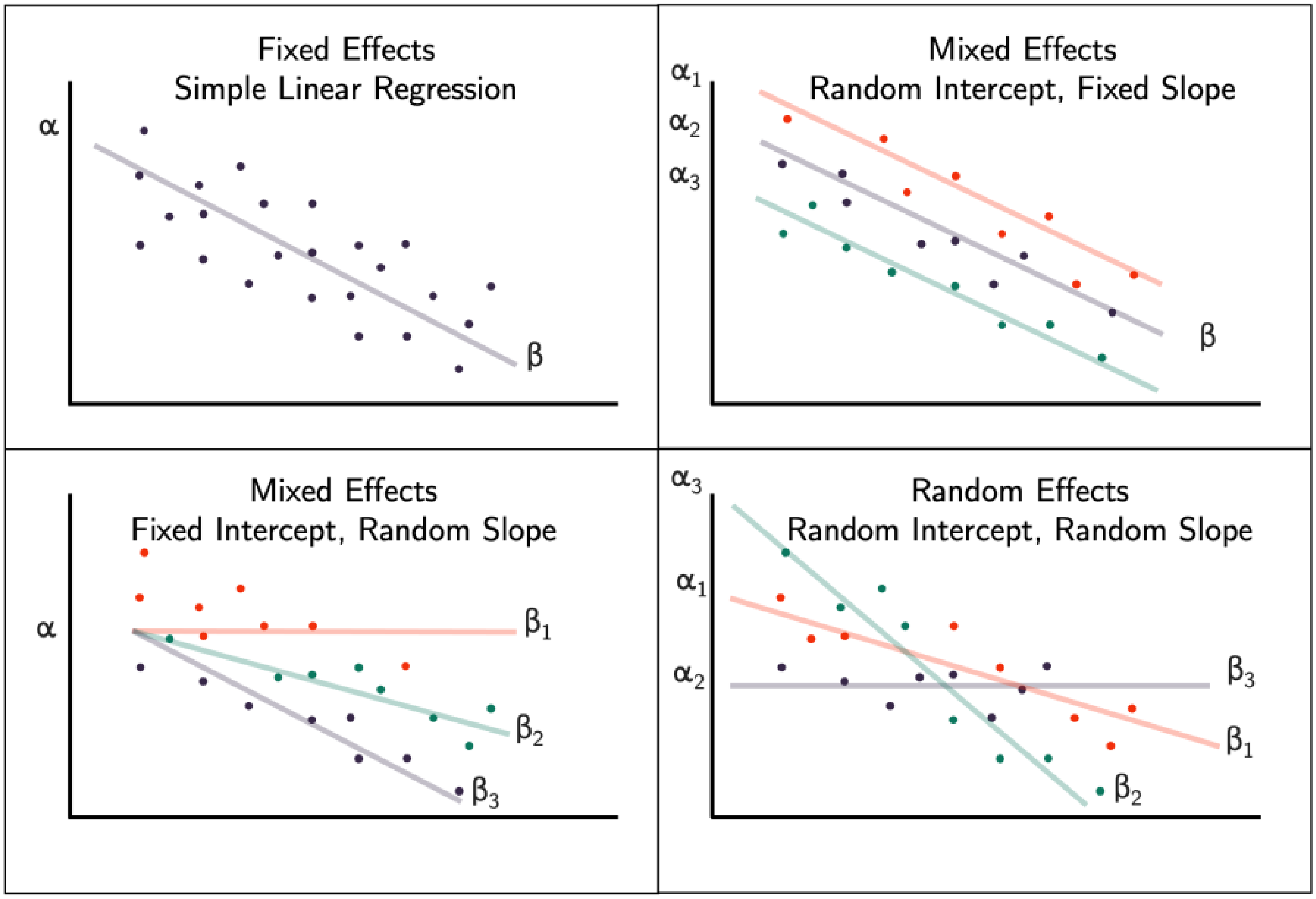
$$Y_{0j} = \gamma_{00} + u_{0j} + (\gamma_{10} + u_{1j})X_{ij} + r_{ij}$$

$$= \underline{\gamma_{00} + \gamma_{10}X_{ij}} + \qquad \underline{u_{0j} + u_{1j}X_{ij} + r_{ij}}$$

- Error Term이 복잡함: Group들 사이의 분산(u_{oj}, u_{1j})과 Individual들 사이의 Group내 분산(r_{ij})이 동시에 존재
- OLS로는 계산이 어렵기에 Maximum Likelihood Estimation (https://www.youtube.com/watch?v=XhlfVtGb19c)을 활용

Fixed Effect and Random Effect (cont'd)

출처 (https://bookdown.org/steve_midway/DAR/random-effects.html)



Module II: Sum-up

- HLM은 서로 다른 수준의 분석단위를 하나의 모델에 포함시켜 하위 수준과 상위 수준의 모수를 동 시에 추정 가능하도록 하는 통계 방법
- 각 개인은 그가 속한 지역이나 집단의 특성으로 영향을 받고 있으며, 특정 조직 또는 집단에 속한 개인들은 그와 다른 집단이나 지역에 속한 개인들과 구별되는 공통 특성 가짐
  - 종속변수: 개인 수준에서 측정
  - 독립변수: 하위수준(개인) & 상위수준(집합단위)
  - 선형 모형이기에 변수들 사이의 선형관계를 가정
- 어떤 변수들이 어떤 수준에 속하는지 지정 가능, 수준 간 교차 및 상호작용효과 고려가능
- 기존 회귀분석 방법에 비해 효과적이며, 잔차의 독립성에서 자유롭고, 분산을 수준별로 산출가능

Module III: HLM의 모형 구체화

Model Specification with R

- R 설치과정 참고: https://www.youtube.com/watch?v=LVkhk4MXQAg&t (https://www.youtube.com/watch?v=LVkhk4MXQAg&t)
- 기초 R 연습 참고 사이트: https://www.youtube.com/watch?v=jLcDVcgQpPI&list=PLKtLBdGREmMmw86INhCWxJNfwBUcnxgZ6 (https://www.youtube.com/watch?v=jLcDVcgQpPI&list=PLKtLBdGREmMmw86INhCWxJNfwBUcnxgZ6)
- 기초 R 문서 버전: R for Data Science (https://r4ds.had.co.nz/)
- 사용하는 패키지: `lmerTest` , `lme4` , `bruceR`

데이터 불러오기

```
# install.packages("pacman")
pacman::p_load("tidyverse", "magrittr", # data분석 필수 패키지
               "broom", # 분석결과 정리해주는 패키지
               "lmerTest", "lme4", "merTools", "nlme", # HLM 패키지 (앞으로 갈 수록 최신)
               "bruceR", # HLM 결과 정리 표 생성
               "readstata13", # statafile (.dta) 열어주는 패키지
               "skimr", "psych") # 요약표, 심리학 분석
```

```


bruceR (version 0.8.9)
BRoadly Useful Convenient and Efficient R functions

Packages also loaded:
√ dplyr         √ emmeans       √ ggplot2
√ tidyr         √ effectsize    √ ggtext
√ stringr       √ performance   √ cowplot
√ forcats       √ lmerTest      √ see
√ data.table

Main functions of `bruceR`:
cc()            Describe()  TTEST()
add()           Freq()      MANOVA()
.mean()         Corr()      EMMEANS()
set.wd()        Alpha()     PROCESS()
import()        EFA()       model_summary()
print_table()   CFA()       lavaan_summary()

https://psychbruce.github.io/bruceR/

These R packages are dependencies of `bruceR` but not installed:
vars, phia
***** Please Install All Dependencies *****
install.packages("bruceR", dep=TRUE)
```

```
## 자신이 설정하고 싶은 곳으로 설정
setwd("E:/OneDrive - SNU/(B) 대학원/세미나/HLM/hlm")
getwd()
```

```
[1] "E:/OneDrive - SNU/(B) 대학원/세미나/HLM/hlm"
```

```
# Read data
data_lv1 <- read.dta13("./HSB1.dta")
data_lv2 <- read.dta13("./HSB2.dta")
```

- 정상적으로 로드 되었는지 확인

```
# Size
dim(data_lv1)
```

```
[1] 7185    5
```

```
dim(data_lv2)
```

```
[1] 160   4
```

Data 설명: High School and Beyond (HS&B)

High School and Beyond (HS&B) is a national *longitudinal* study originally funded by the United States Department of Education's National Center for Education Statistics (NCES) as a part of their longitudinal studies program.

Purpose was to document the **educational, vocational, and personal development of young people** following them over time as they begin to take on adult roles and responsibilities

- **Level-1 file**: HSB1.dta, 7,185 observations with 4 variables
    - MINORITY: an indicator for student ethnicity (1 = minority, 0 = other)
    - FEMALE: an indicator for student gender (1 = female, 0 = male)
    - SES: a standardized scale constructed from variables measuring parental education, occupation, and income
    - MATHACH: a measure of mathematics achievement
- **Level-2 file**: HSB2.dta, 160 schools with 3 variables
    - SIZE: school enrollment
    - SECTOR (1 = Catholic, 0 = public)
    - HIMNTY (1 = more than 40% minority enrollment, 0 = less than 40%)

Data Glimpse

- `data_lv1` glimpse

```
glimpse(data_lv1)
```

```
Rows: 7,185
Columns: 5
$ id       <chr> "1224", "1224", "1224", "1224", "1224", "1224", "1224", "1224", "1224", "1224", "1224",
"1224", "1…
$ minority <dbl> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0,
0, 0, 0,…
$ female   <dbl> 1, 1, 0, 0, 0, 0, 1, 0, 1, 0, 1, 1, 0, 1, 0, 1, 0, 0, 1, 0, 1, 1, 1, 0, 1, 0, 1, 1, 1, 1,
1, 1, 1,…
$ ses      <dbl> -1.528, -0.588, -0.528, -0.668, -0.158, 0.022, -0.618, -0.998, -0.888, -0.458, -1.448, -
0.658, -0.…
$ mathach  <dbl> 5.876, 19.708, 20.349, 8.781, 17.898, 4.583, -2.832, 0.523, 1.527, 21.521, 9.475, 16.057,
21.178, …
```

- `data_lv2` glimpse

```
glimpse(data_lv2)
```

```
Rows: 160
Columns: 4
$ id      <chr> "1224", "1288", "1296", "1308", "1317", "1358", "1374", "1433", "1436", "1461", "1462", "1
477", "14…
$ size    <dbl> 842, 1855, 1719, 716, 455, 1430, 2400, 899, 185, 1672, 530, 531, 1921, 100, 400, 1357, 95
9, 1646, 1…
$ sector  <dbl> 0, 0, 0, 1, 1, 0, 0, 1, 1, 0, 1, 1, 0, 0, 1, 0, 0, 0, 0, 1, 1, 1, 0, 1, 0, 1, 0, 1, 0, 0,
0, 1, 1, …
$ himinty <dbl> 0, 0, 1, 0, 1, 0, 0, 0, 0, 0, 1, 0, 1, 1, 0, 0, 0, 0, 0, 0, 1, 1, 0, 1, 0, 0, 0, 0, 1, 0,
1, 0, 0, …
```

Data Summarise

- `data_lv1` 기술통계

	vars	n	mean	sd	median	trimmed	mad	min	max	range	skew	kurtosis	se
id*	1	7185	79.96	45.44	79.00	79.71	57.82	1.00	160.00	159.00	0.04	-1.17	0.54
minority	2	7185	0.27	0.45	0.00	0.22	0.00	0.00	1.00	1.00	1.01	-0.98	0.01
female	3	7185	0.53	0.50	1.00	0.54	0.00	0.00	1.00	1.00	-0.11	-1.99	0.01
ses	4	7185	0.00	0.78	0.00	0.02	0.85	-3.76	2.69	6.45	-0.23	-0.38	0.01
mathach	5	7185	12.75	6.88	13.13	12.91	8.12	-2.83	24.99	27.82	-0.18	-0.92	0.08

- `data_lv2` 기술통계

	vars	n	mean	sd	median	trimmed	mad	min	max	range	skew	kurtosis	se
id*	1	160	80.50	46.33	80.5	80.50	59.30	1	160	159	0.00	-1.22	3.66
size	2	160	1097.83	629.51	1061.0	1058.24	695.34	100	2713	2613	0.46	-0.61	49.77
sector	3	160	0.44	0.50	0.0	0.42	0.00	0	1	1	0.25	-1.95	0.04
himinty	4	160	0.28	0.45	0.0	0.22	0.00	0	1	1	1.00	-1.01	0.04

Data Merge: `dplyr` package - joins

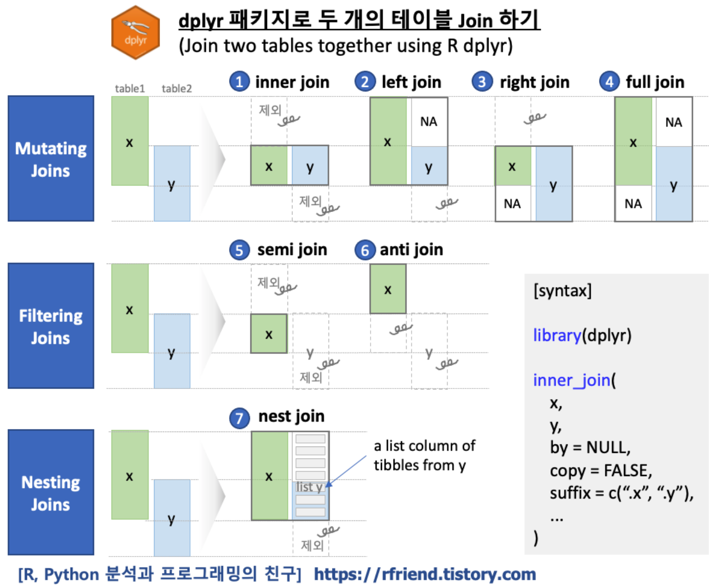
`dplyr` 패키지는 다양한 데이터 merge 함수들을 제공하고 있음. 이중 HLM에서 자주 활용되는 `left_join`, `right_join`, `full_join`, `inner_join`, `semi_join`, `anti_join` 등에 대해 간단히 다루고 넘어감.

- `dplyr` join 함수들의 유형형
  - `left_join` (`right_join`): Join matching rows from y to x (x to y)
  - `full_join` : Join data. Retain all values, all rows
  - `inner_join` : Join data. Retain only rows in both sets
  - `semi_join` : All rows in a that have a match in b
  - `anti_join` : All rows in a that do not have a match in b

- Join 함수의 구조

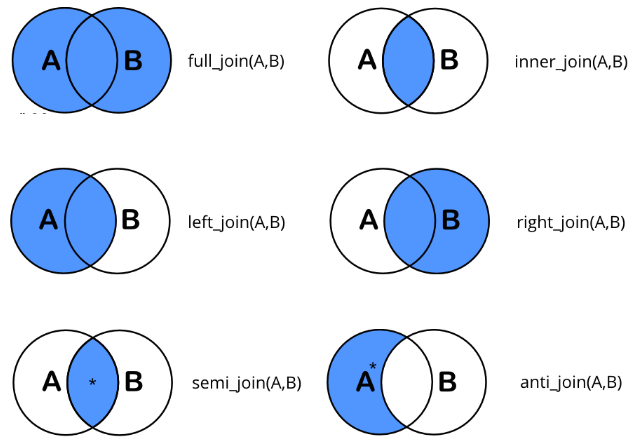
```
left_join(
  x, # Level 1 data-set name
  y, # Level 2 data-set name
  by = c("id_x" = "id_y"), # 각각의 데이터 셋에서 어떤 변수를 기준으로 merge되는지 설정
  copy = FALSE, # 가만히 두기
  suffix = c(".x", ".y"), # 만약 id 이외에 서로 겹치는 변수가 있을때 어느 데이터셋인지
  ...,
  keep = FALSE # 가만히 두기
)
```

Data Merge: `dplyr` package - joins (Cont'd)



Join 함수들

출처 (https://rfriend.tistory.com/625)



벤다이어그램

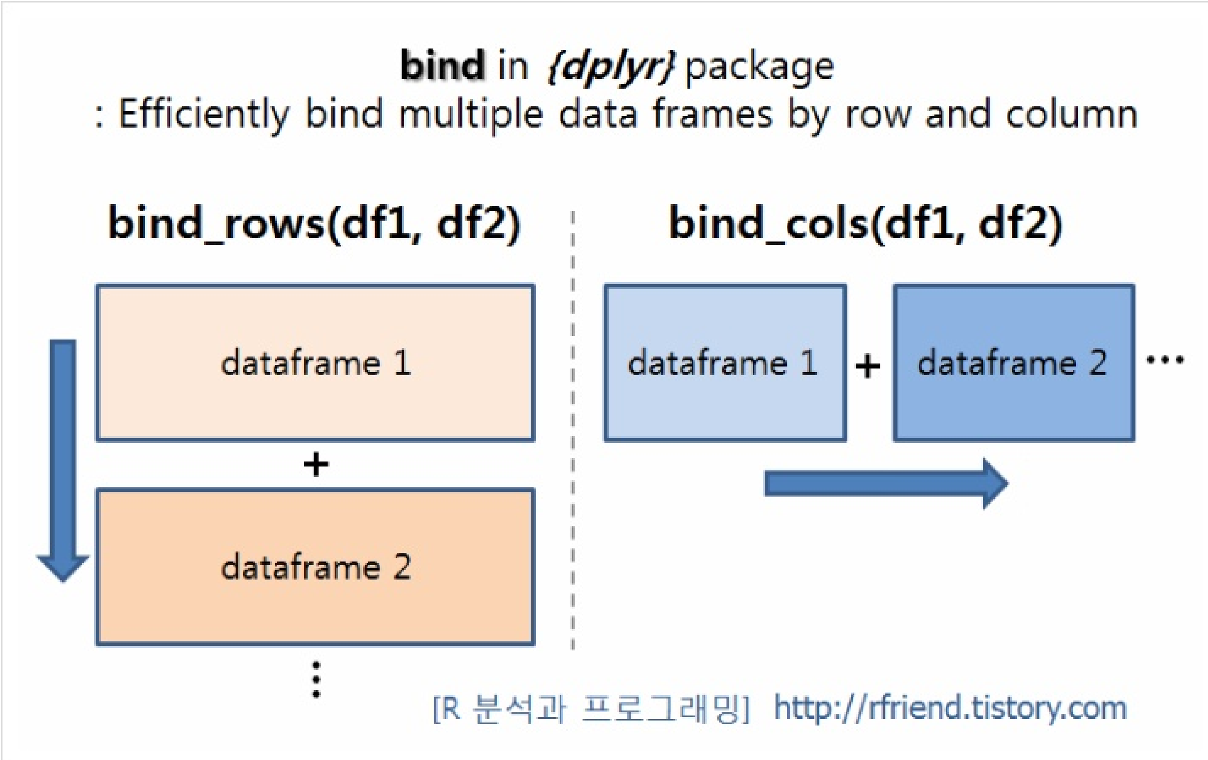
출처 (http://www2.stat.duke.edu/~cr173/Sta523_Fa15/gis.html)

Data Merge: `dplyr` package - binds

- `bind_rows()`
  - 다수의 데이터 프레임(티블)을 행 기준으로 합치기 (binding multiple data frames by row)
- `bind_cols()`
  - 다수의 데이터 프레임(티블)을 열 기준으로 합치기 (binding multiple data frames by columns)
- bind 함수의 구조

```
# row 병합
bind_rows(list(data*, data*, ... ))

## id의 경우 어떤 data frame, tibble에서 결합되었는지 확인 용도
bind_rows(list(a = one, b = two), .id = "id")

# column 병합
bind_cols(list(data*, data*, ... ), .name_repair = c("unique", "universal", "minimal"))
```

Data Merge: `dplyr` package - binds (Cont'd)



출처 (https://rfriend.tistory.com/248)

Data Merge: `data_lv1` and `data_lv2`

Pipe Operator: `%>%`

- lhs `%>%` rhs는 lhs의 결과를 rhs의 첫번째 변수로 넘겨주는 역할. `'.'` 은 앞의 값의 위치를 구체적으로 지정하기 위해 사용
- x `%>%` f is equivalent to f(x)
- x `%>%` f(y) is equivalent to f(x, y)
- x `%>%` f(y, .) is equivalent to f(y, x)

data_lv1와 data_lv2 데이터 병합

```
data_merged <- data_lv1 %>%
  left_join(data_lv2, by = "id")
head(data_merged, 5)
```

```
    id minority female    ses mathach size sector himinty
1 1224        0      1 -1.528   5.876  842      0       0
2 1224        0      1 -0.588  19.708  842      0       0
3 1224        0      0 -0.528  20.349  842      0       0
4 1224        0      0 -0.668   8.781  842      0       0
5 1224        0      0 -0.158  17.898  842      0       0
```

id	minority	female	ses	mathach	size	sector	himinty	ses_grandmc	meanses	ses_groupmc
1224	0	1	-1.53	5.88	842	Public	0	-1.53	-0.43	-1.09
1224	0	1	-0.59	19.71	842	Public	0	-0.59	-0.43	-0.15
1224	0	0	-0.53	20.35	842	Public	0	-0.53	-0.43	-0.09

id	minority	female	ses	mathach	size	sector	himinty	ses_grandmc	meanses	ses_groupmc
1224	0	0	-0.67	8.78	842	Public	0	-0.67	-0.43	-0.23
1224	0	0	-0.16	17.90	842	Public	0	-0.16	-0.43	0.28

Pooled-OLS and Between Model

```
다음의 패키지를 부착합니다: 'patchwork'
```

```
The following object is masked from 'package:cowplot':

    align_plots
```

```
The following object is masked from 'package:MASS':

    area
```

```
Warning: Using `size` aesthetic for lines was deprecated in ggplot2 3.4.0.
ℹ Please use `linewidth` instead.
```

```
`geom_smooth()` using formula = 'y ~ x'
```

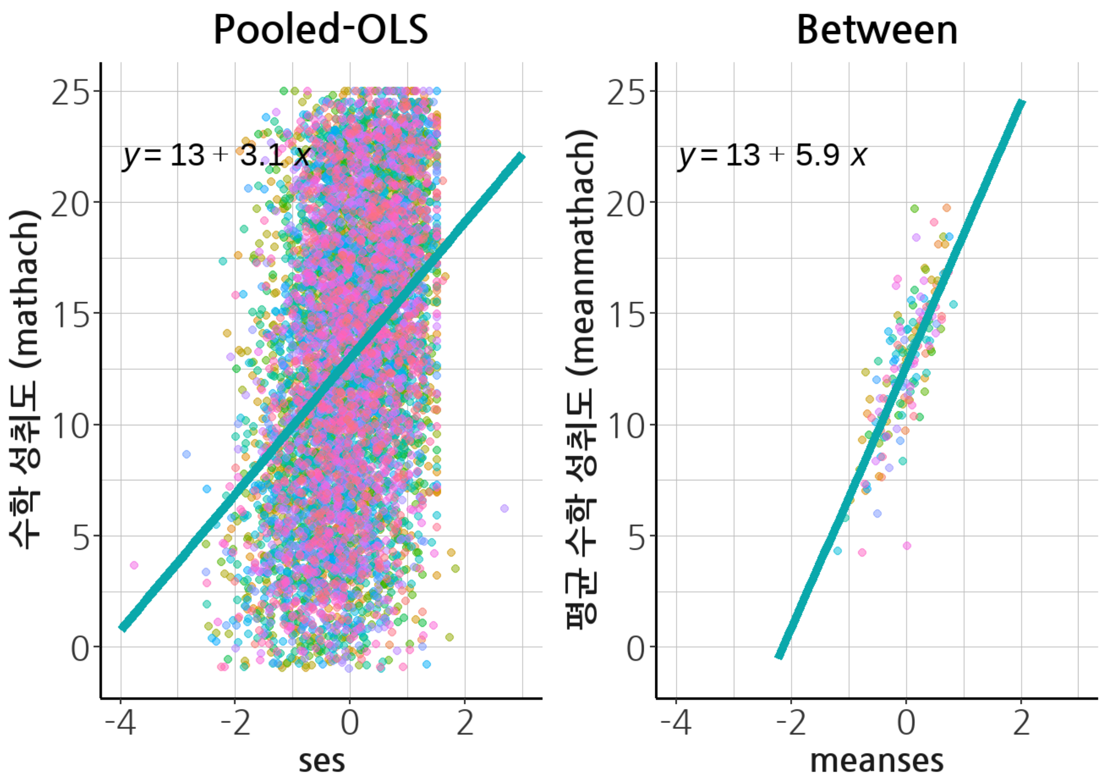
```
Warning: Removed 135 rows containing non-finite values (`stat_smooth()`).
```

```
Warning: Removed 135 rows containing non-finite values (`stat_regline_equation()`).
```

```
Warning: Removed 135 rows containing missing values (`geom_point()`).
```

```
`geom_smooth()` using formula = 'y ~ x'
```

```
Warning: Removed 31 rows containing missing values (`geom_smooth()`).
```



Five models in HLM (Bryk & Raudenbush, 1992)

Overview of HLM Two-Level Models

	(1) One-way ANOVA	**(2) Means-as-Outcomes (Between)**	**(3) One-way ANCOVA**	**(4) Random Coefficient**	**(5) Intercept-and-Slopes-as-Outcomes**
Level-1-Models:			(Different Intcpt, Same Slope)	(Different Intcpt, Different Slope)	**<우리의 목표>** (Different Intcpt, Different Slope)
For Level-1 Intercept	$Y_{ij} = \beta_{0j} + r_{ij}$	$Y_{ij} = \beta_{0j} + r_{ij}$	$Y_{ij} = \beta_{0j} + \beta_{1j}X_{ij} + r_{ij}$	$Y_{ij} = \beta_{0j} + \beta_{1j}X_{ij} + r_{ij}$	$Y_{ij} = \beta_{0j} + \beta_{1j}X_{ij} + r_{ij}$
Level 1 Independent Variable	NO	NO	YES	YES	YES
Level-2-Models:					
For Level-1 Intercept:	$\beta_{0j} = \gamma_{00} + u_{0j}$	$\beta_{0j} = \gamma_{00} + \gamma_{01}W_j + u_{0j}$	$\beta_{0j} = \gamma_{00} + u_{0j}$	$\beta_{0j} = \gamma_{00} + u_{0j}$	$\beta_{0j} = \gamma_{00} + \gamma_{01}W_j + u_{0j}$
Level 2 Independent Variable	NO	**YES**	NO	NO	YES
For Level 1 Slopes	NO	NO	$\beta_{1j} = \gamma_{10}$	$\beta_{1j} = \gamma_{10} + u_{ij}$	$\beta_{1j} = \gamma_{10} + \gamma_{11}W_j + u_{ij}$
Level 2 Independent Variables			Fixed	Random	Yes (Sometimes)

변수들에 대한 설명

- 분석 목적: Level-1 변수인 SES (Social Economic Status)와 Level-2 변수인 Sector (Public or Private)가 학생들의 수학성취도에 미치는 영향
- 독립변수: `ses` , `sector`
- 종속변수: `mathach`
- i는 개인수준 (Level-1)의 첨자, j는 집단수준 (Level-2)의 첨자

- Y_{ij}: j 번째 학교에 다니는 i번째 학생의 수학성취도 점수
- β_{0j}는 j번째 학교의 평균 수학성취도 점수

- 평균 학급당 인원: 44.91명 (sd: 11.85)

Model 0. Preliminary Analysis

One-Way ANOVA 집단수준의 Mean Squares들의 값과 개인수준의 Mean Squares들의 값을 비교함을 통해서, 개인수준의 변량대비 집단수준의 변량 비교

→ 쉽게 말해 **집단간 모평균(여기서는 수학성취도)이 서로 다른가**? 수준효과가 존재하는가?

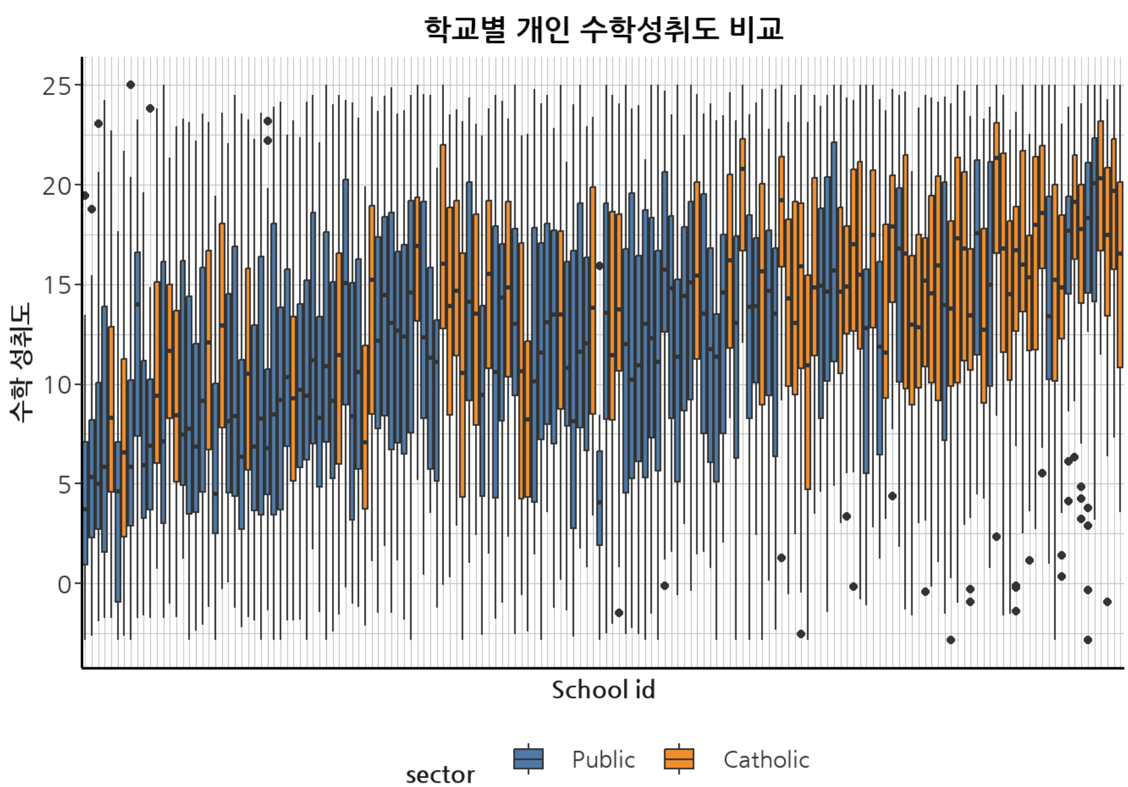
$$SST = SSA + SSE \quad \sum_{i=1}^{a}\sum_{j=1}^{r}(Y_{ij}-\bar{Y})^2 = \sum_{i=1}^{a} r(\bar{Y_{i.}}-\bar{Y})^2 + \sum_{i=1}^{a}\sum_{j=1}^{r}(Y_{ij}-\bar{Y_{i.}})^2$$

- SSA (sum of squares of treatment): 집단변량 제곱합 - 집단수준 변동
- SSE (sum of squares of error): 오차제곱합 - 집단 내 변동
- **집단간 평균 차이가 존재할 때 HLM의 가장 최소한의 근거가 됨**

```
model1 <-  aov(mathach~id, data=data_merged)
anova(model1)
```

term	df	sumsq	meansq	statistic	p.value
id	159	64906.96	408.220	10.429	0
Residuals	7025	274969.98	39.142		

Model 0. Preliminary Analysis: Visualization



Model 1. One-way ANOVA

- Level 1: $Y_{ij} = \beta_{0j} + \epsilon_{ij} \quad \epsilon_{ij} \sim N(0, \sigma^2)$
- Level 2: $\beta_{0j} = \gamma_{00} + u_{0j}, \quad u_{0j} \sim N(0, \tau_{00})$
- 급간상관계수와 신뢰도 계산이 주 목적으로 모형에 아무런 predictor variable들이 투입되지 않음

```
model1 <- lmer(mathach ~ 1 + (1 | id), data=data_merged)
summary(model1)
```

```
Linear mixed model fit by REML. t-tests use Satterthwaite's method ['lmerModLmerTest']
Formula: mathach ~ 1 + (1 | id)
   Data: data_merged

REML criterion at convergence: 47116.8

Scaled residuals:
    Min      1Q  Median      3Q     Max
-3.0631 -0.7539  0.0267  0.7606  2.7426

Random effects:
 Groups   Name        Variance Std.Dev.
 id       (Intercept)  8.614   2.935
 Residual             39.148   6.257
Number of obs: 7185, groups:  id, 160

Fixed effects:
            Estimate Std. Error       df t value Pr(>|t|)
(Intercept)  12.6370     0.2444 156.6473   51.71   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Model 1. One-way ANOVA: Model Fit

```
HLM_summary(model1,test.rand = T, digits = 3)
```

```


Hierarchical Linear Model (HLM)
(also known as) Linear Mixed Model (LMM)
(also known as) Multilevel Linear Model (MLM)

Model Information:
Formula: mathach ~ 1 + (1 | id)
Level-1 Observations: N = 7185
Level-2 Groups/Clusters: id, 160

Model Fit:
AIC = 47122.793
BIC = 47143.433
R_(m)² = 0.00000  (Marginal R²: fixed effects)
R_(c)² = 0.18035  (Conditional R²: fixed + random e
ffects)
Omega² = 0.18903  (= 1 - proportion of unexplained
variance)

Fixed Effects:
Unstandardized Coefficients (b or γ):
Outcome Variable: mathach
────────────────────────────────────────────────────
───────────────
                b/γ    S.E.     t    df     p
[95% CI of b/γ]
────────────────────────────────────────────────────
───────────────
(Intercept)  12.637 (0.244) 51.71 156.6 <.001 ***
[12.154, 13.120]
────────────────────────────────────────────────────
───────────────
'df' is estimated by Satterthwaite approximation.

Random Effects:
──────────────────────────────────────────
 Cluster  K   Parameter   Variance     ICC
──────────────────────────────────────────
 id       160 (Intercept)  8.61402 0.18035
 Residual                 39.14832
──────────────────────────────────────────

ANOVA-like table for random-effects: Single term de
letions

Model:
mathach ~ (1 | id)
         npar logLik   AIC    LRT Df Pr(>Chisq)
<none>      3 -23558 47123
(1 | id)    2 -24052 48107 986.12  1  < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05
'.' 0.1 ' ' 1
```

```
Model Information:
Formula: mathach ~ 1 + (1 ¦ id)
Level-1 Observations: N = 7185
Level-2 Groups/Clusters: id, 160

Model Fit:
```

통계 모델간의 적합성 비교 기준

- k=투입되는 변수의 갯수, n=데이터의 갯수
- 다음 세가지 기준은 0과 가까워질 수록 해당 모형의 Model Fit이 좋아짐
  - $Deviance$ = -2 * ln(Likelihood)
  - AIC = -2 * ln(Likelihood) + 2 * k
  - BIC = -2 * ln(Likelihood) + k * log(n)

  변수가 많은 모형일수록 우도가 0과 가까워지기에, AIC와 BIC는 Overfitting 문제해결과 모형 Parsimony를 위해 독립변수 증가에 대한 패널티 부여. LRtest로 모형 간 차이를 검정하여 변수 투입으로 인한 model fit 개선효과를 확인.
- 다음 기준들은 클 수록 좋음
  - $R^2(m)$ = Pooled OLS의 R^2 과 동일
  - $R^2(c)$ = Pooled OLS의 R^2 에 random effect의 효과
  - $Omega^2$ = $\frac{SS_{effect}-(df_{effect})(MS_{error})}{MS_{error}+SS_{total}}$

Model 1. One-way ANOVA: 계수해석

```
AIC = 47122.793
BIC = 47143.433
R_(m)² = 0.00000  (Marginal R²: fixed effects)
R_(c)² = 0.18035  (Conditional R²: fixed + random effects)
Omega² = 0.18903  (= 1 - proportion of unexplained variance)
```

```
HLM_summary(model1,test.rand = T, digits = 3)
```

```
Hierarchical Linear Model (HLM)
(also known as) Linear Mixed Model (LMM)
(also known as) Multilevel Linear Model (MLM)

Model Information:
Formula: mathach ~ 1 + (1 | id)
Level-1 Observations: N = 7185
Level-2 Groups/Clusters: id, 160

Model Fit:
AIC = 47122.793
BIC = 47143.433
R_(m)² = 0.00000  (Marginal R²: fixed effects)
R_(c)² = 0.18035  (Conditional R²: fixed + random e
ffects)
Omega² = 0.18903  (= 1 - proportion of unexplained
variance)

Fixed Effects:
Unstandardized Coefficients (b or γ):
Outcome Variable: mathach
────────────────────────────────────────────────────
───────────────
                b/γ    S.E.     t    df     p
[95% CI of b/γ]
────────────────────────────────────────────────────
───────────────
(Intercept)  12.637 (0.244) 51.71 156.6 <.001 ***
[12.154, 13.120]
────────────────────────────────────────────────────
───────────────
'df' is estimated by Satterthwaite approximation.

Random Effects:
───────────────────────────────────────────
 Cluster   K   Parameter   Variance     ICC
───────────────────────────────────────────
 id       160 (Intercept)   8.61402 0.18035
 Residual                  39.14832
───────────────────────────────────────────

ANOVA-like table for random-effects: Single term de
letions

Model:
mathach ~ (1 | id)
         npar logLik   AIC    LRT Df Pr(>Chisq)
<none>      3 -23558 47123
(1 | id)    2 -24052 48107 986.12  1  < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05
'.' 0.1 ' ' 1
```

```
Fixed Effects:
Unstandardized Coefficients (b or γ):
Outcome Variable: mathach
────────────────────────────────────────────────────
            b/γ    S.E.     t    df     p     [95% CI of b/γ]
────────────────────────────────────────────────────
```

추정된 고정효과 모수: 평균 수학성취도

- γ_{00} = 12.637

추정된 임의효과 모수:

- Level 1 (개인수준): 평균적으로 동일 학교 내에서 학생들의 수학성취도가 어느정도 차이가 있는가?

  $\hat{var}(e_{ij}) = \hat{\sigma}^2 = 39.14832$

- Level 2 (집단수준): 학교 간 수학성취도의 평균이 서로 얼마나 다른지?

  $\hat{var}(u_{0j}) = \hat{var}(\beta_{0j}|\gamma_{00}) = \hat{\tau} = 8.61402$

집단수준 분산 (τ; 학교간 차이) 검정 (Significance test for the intercept variance):

- Absolute Null Model (τ_{00})만 투입된 모형인 Pooled-OLS모형과 HLM 모형의 비교
- χ^2 검정 (Likelihood-Ratio Test)을 통해 Null Model 대비 `p-value` 값이 2.2e-16로 매우 작게 통계적으로 유의한 것으로 나타남

Model 1. One-way ANOVA: ICC

```
(Intercept)  12.637 (0.244) 51.71 156.6 <.001 *** [12.154, 13.120]
───────────────────────────────────────────────────────────────────
'df' is estimated by Satterthwaite approximation.

Random Effects:
────────────────────────────────────────────
 Cluster  K   Parameter   Variance     ICC
────────────────────────────────────────────
 id       160 (Intercept)  8.61402 0.18035
 Residual                 39.14832
────────────────────────────────────────────

ANOVA-like table for random-effects: Single term deletions

Model:
mathach ~ (1 | id)
         npar logLik   AIC    LRT Df Pr(>Chisq)
<none>      3 -23558 47123
(1 | id)    2 -24052 48107 986.12  1  < 2.2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

```
HLM_ICC_rWG(data_merged, group="id", icc.var="mathach")
```

```
------ Sample Size Information ------

Level 1: N = 7185 observations ("mathach")
Level 2: K = 160 groups ("id")

       n (group sizes)
Min.          14.00000
Median        47.00000
Mean          44.90625
Max.          67.00000

------ ICC(1), ICC(2), and rWG ------

ICC variable: "mathach"

ICC(1) = 0.180 (non-independence of data)
ICC(2) = 0.901 (reliability of group means)

rWG variable: "mathach"

rWG (within-group agreement for single-item measures)
─────────────────────────────────────────────
      Min. 1st Qu. Median  Mean 3rd Qu.  Max.
─────────────────────────────────────────────
rWG  0.000   0.270  0.381 0.386   0.515 0.806
─────────────────────────────────────────────
```

Model 1. One-way ANOVA: ICC (Cont’d)

Intraclass correlation (ICC)(1):

- ICC(1)은 전체 관찰 분산에서 집단 간 분산이 차지하는 비율 (강상진, 2016)
- ICC(1) is typically interpreted as a measure of effect size (Bliese, 2000; Bryk & Raudenbush, 1992)
- 전체 분산(Level-2 분산과 Level-1 또는 잔차 분산의 합)에 대한 Level-2 분산(집단 평균의 분산)의 비율이 클수록, 집단 간(between)의 유사성보다 집단 내부(within)간의 유사성이 큼을 의미

$$ICC(1) = \frac{\text{학교 간 분산}}{\text{전체 관찰분산}} = \frac{Var(\beta_{0j})}{Var(Y_{ij})} = \frac{\tau}{\sigma^2 + \tau} = \frac{8.614}{39.148 + 8.614} = 0.18035$$

- 만약 ICC(1)의 값이 0이라면 한 집단에 속한 응답치들 간의 유사성이 다른집단에 속한 응답치들과 보이는 유사성과 다르지 않음 (일반적으로 0.05~0.25 정도)
- The value of .01 might be considered a small effect, a value of .10 might be considered a medium effect, and a value of .25 might be considered a large effect (Murphy & Myors, 1998 (https://psycnet.apa.org/record/2014-24220-000)).
- 그러나 통일된 기준은 존재하지 않으며, 이론적으로 집단을 고려함을 통해 설명되는 정도가 어느정도인지를 이해하는 것이 중요함. ICC 값이 매우 작아 0에 가깝더라도 측정값과 다른 측정값 사이의 관계가 모든 집단에서 동일하다는 것을 의미하지 않음 (Nezlek, 2008 (https://doi.org/10.1111/j.1751-9004.2007.00059.x))
- Simulated situation only 1% of the variance is attributed to group membership ICC(1)=.01 and, still, strong group-level relationships were detected (Bilese, 1998 (https://doi.org/10.1177/109442819814001))

Model 1. One-way ANOVA: ICC (Cont'd)

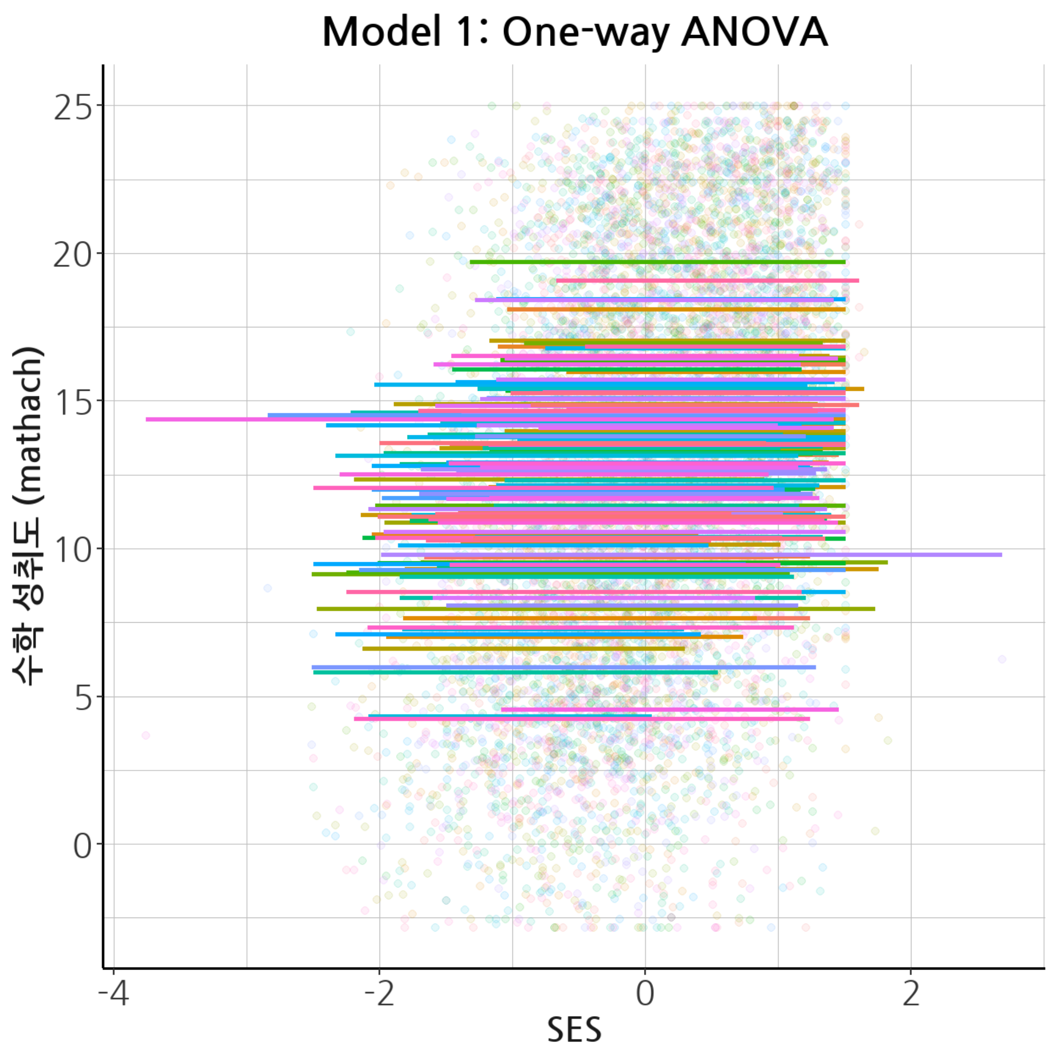
Intraclass correlation (ICC)(2) :

$$ICC(2) = \frac{Var(\text{진 점수})}{Var(Y)} = \frac{\tau_{00}}{\tau_{00} + \sigma^2/n_j}$$

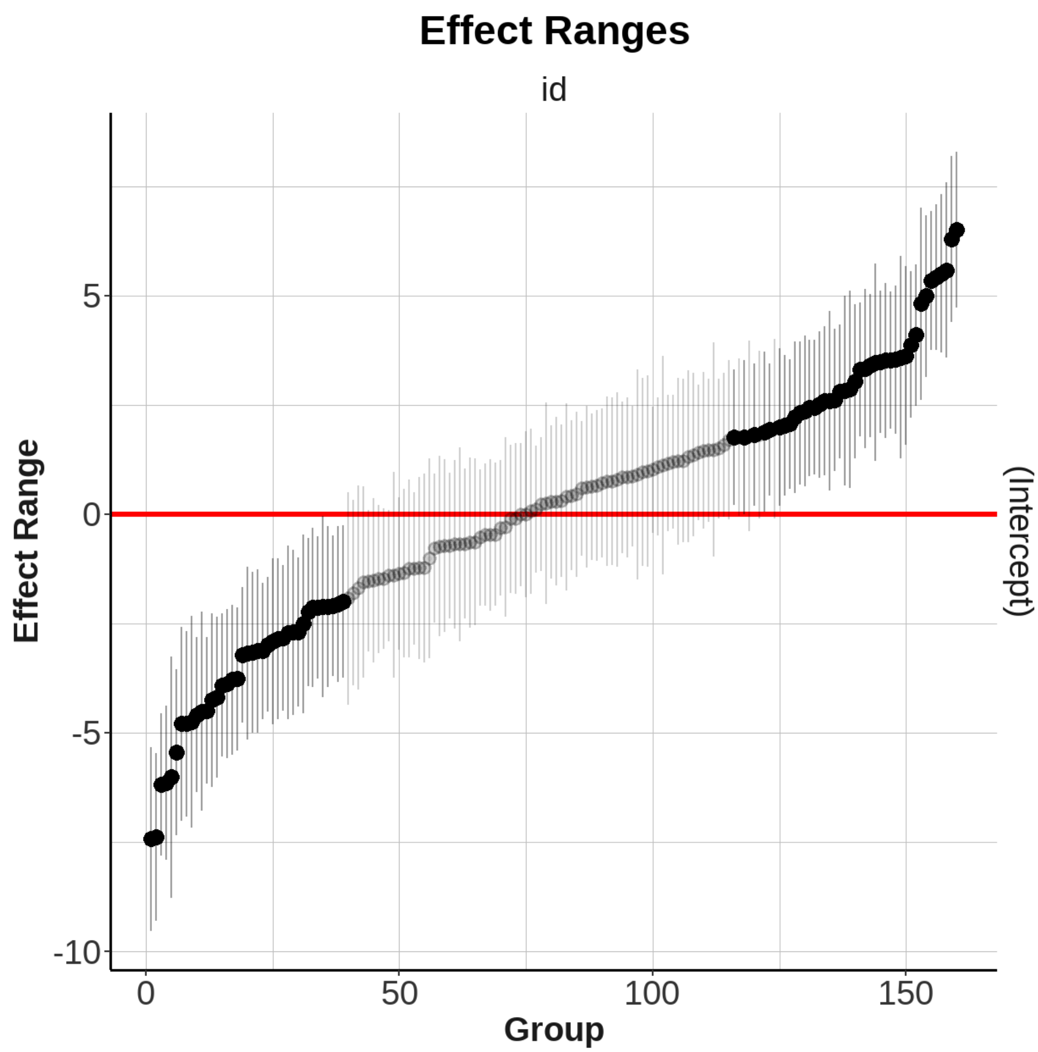
- ICC(2)는 집단간 평균의 신뢰도를 측정하기 위함이며, 집단별 평균(학교별 수학성취도 평균)은 β_{0j} 표본에 따라 그 값이 달라지기 때문에 통계적 추정의 차원에서는 의미가 없으나, 잔차분석에서 제공하는 β_{0j}가 어느정도 신뢰로운 값인지 알려줌. β_{0j}가 높으면 학교 정보로서의 가치가 높고, 낮으면 β_{0j}에 의한 평가가 위험함
- ICC(2) <0.40 are poor, those from 0.40 to 0.75 are fair to good, and those >0.75 are excellent (Fleiss, 1986 (https://doi.org/10.1002/bimj.4710300308))
- τ_{00}이 크거나 각 학교의 표본이 크면 이 값은 커지게 됨
- 일반적으로 무선효과의 추정은 Random Level-1 coefficients에 대해 이 Level-1 모형의 $Y_{\cdot j}$와 Level-2 모형의 $\hat{\gamma_{00}}$ 를 동시에 고려하는 추정치 (a weighted combination (WLS), known as a Bayes estimator)
- HLM 은 $Y_{\cdot j}$ 의 신뢰도가 높으면 $Y_{\cdot j}$ 가 더 많이 가중되고 그 신뢰도가 낮으면 Level-2 모형에서 얻어지는 $\hat{\gamma_{00}}$ 값에 더 많은 가중치를 주는 방식으로 β^*_{0j} 를 추정한다. 이러한 이유로 β^*_{0j}은 전체 평균 (grand mean) γ_{00}로 집약되는 모습을 보여서 shrinkage estimator 라고 불린다.

Model 1. One-way ANOVA: Visualization

Model 1. One-way ANOVA: Visualization



Model 1. One-way ANOVA: Visualization



Model 3. One-Way ANCOVA

- Model 3. One-Way ANCOVA는 Random intercept model로 불리기도 함.
- ANOVA의 경우, average or expected outcomes among groups (level-2 units) 이외에 관심이 없을 때는 유용함. 그러나, Level 1 독립변수들이 종속변수와 Level 2에 상관되어 있는데 Level 1의 독립변수를 모형에 포함하지 않으면 L-2 의 효과를 편파적으로 추정 $\rightarrow$ ANCOVA필요
- 예를 들어, 교육연한이나 일 경험은 임금수준에 긍정적 효과를 주는 것으로 알려져 있는데 특정 업무를 하는 사람들 중에서 남자들이 교육연한이나 일경험에 대한 변수 값이 더 높을 때 그 업무에 종사하는 여자와 남자의 평균임금수준에 대한 비교분석 결과는 교육연한이나 일 경험을 통제했는지 여부에 달려있게 될 것
- One-Way ANCOVA는 level 1 독립변수를 투입하고, 기울기를 fixed하게 고려
  - Level 1: $Y_{ij} = \beta_{0j} + \beta_{1j}X_{ij} + \epsilon_{ij} \quad \epsilon_{ij} \sim N(0, \sigma^2)$
  - Level 2 Intercept: $\beta_{0i} = \gamma_{00} + u_{0j} \quad u_{0j} \sim N(0, \tau_{00})$
  - Level 2 Slope: $\beta_{1j} = \gamma_{10}$
- Fixed Effect: γ_{00} = average outcome for sample of groups, γ_{10} = average individual effect (slope) on outcome
- Random Effect: ϵ_{ij} = residuals, u_{0j} = unique effect of group j on average outcome

Model 3. One-Way ANCOVA (Cont’d)

- level-1 predictor(s) 도입하려면, 아래와 같은 문제에 대해 결정해야 함.

1. Whether to introduce random coefficients
2. Whether to center or transform the level-1 predictor (Module III 참고)

- 일반적으로 ANCOVA에서는 Grand Mean Centering을 적용 Covariate 는 Level-1 종속변수와 영향을 주지만 Level-2 를 구성하는 집단(학교)에 따라 개인들의 Covariate 값이 다를 수 있기 때문에

- Grand-mean Centering 후 모형 추정

```
data_merged <- data_merged %>%
  mutate(ses_grandmc = ses - mean(ses))

model3 <- lmer(mathach ~ ses_grandmc + (1 | id), data=data_merged)
icc(model3)
```

```
# Intraclass Correlation Coefficient

    Adjusted ICC: 0.114
  Unadjusted ICC: 0.105
```

```
# the individual random effects (level 2 residuals of the intercept, i.e. the u0i)
ranef(model3)
```

```
$id
     (Intercept)
1224 -1.633782013
1288  0.429297757
1296 -3.446764277
1308  1.682570394
1317 -0.262699062
1358 -1.115449875
1374 -2.269307587
1433  4.386431700
1436  3.492215299
1461  2.077046149
1462 -0.494527543
1477  1.057127724
1499 -3.387101033
1637 -3.065473199
1906  1.834038253
1909  0.918199890
1942  3.015633281
1946  0.201202122
2030 -1.170720641
2208  1.536796934
2277 -1.364910197
2305 -0.016735981
2336  2.407002600
2458  0.689754430
2467 -1.498078883
2526  3.180720523
2626  0.742373212
2629  2.269984840
2639 -3.153634117
2651 -1.431318702
2655  1.187423056
2658 -0.263766660
2755  2.136853529
2768 -1.253125025
2771 -0.001759821
2818  0.806477745
2917 -2.346712972
2990  3.637037409
2995 -2.044971089
3013  0.051510943
3020  1.049554262
3039  2.536959441
3088 -1.986393117
3152  0.415333743
3332  0.132129289
3351 -1.644284322
3377 -1.766764445
3427  5.776697660
3498  1.880887121
3499 -0.378854593
3533 -1.658664326
3610  2.149489940
3657 -1.375208787
3688  0.873002177
3705 -2.461786462
3716 -1.078918051
3838  2.673431076
3881 -0.851891595
3967 -0.073373963
```

3992	0.971925839
3999	-1.279743254
4042	0.621640977
4173	-0.007910967
4223	1.867491463
4253	-2.073726296
4292	1.222765455
4325	0.610489542
4350	-0.944510934
4383	-1.042789018
4410	0.491915785
4420	1.410357353
4458	-3.733540270
4511	0.888634487
4523	-3.584858544
4530	-1.936373439
4642	1.477495744
4868	-0.985351008
4931	0.230604460
5192	-1.640600068
5404	0.691417857
5619	1.569394118
5640	0.813830169
5650	1.332429256
5667	0.105683246
5720	1.349253252
5761	-0.650262632
5762	-4.528279608
5783	0.062530808
5815	-2.869340685
5819	-0.824553691
5838	0.526191956
5937	2.021213648
6074	1.565909711
6089	2.191001609
6144	-2.597445484
6170	1.638821576
6291	-1.066496058
6366	1.991802777
6397	0.580899695
6415	-0.305717858
6443	-1.857342114
6464	-3.115114547
6469	3.519918742
6484	0.570782645
6578	0.501649614
6600	-0.725062956
6808	-2.670134071
6816	0.540264658
6897	1.385103984
6990	-4.805919757
7011	0.996458333
7101	-0.645285527
7172	-3.313102249
7232	0.087481125
7276	-0.143963102
7332	1.092031481
7341	-2.144525685
7342	-0.370990169
7345	-1.228071613
7364	1.468949771
7635	1.585774418

```
7688  4.651485327
7697  1.969096473
7734 -0.721315752
7890 -2.661658096
7919  0.907960273
8009  0.167188700
8150  1.244182921
8165  2.637283829
8175 -0.444520833
8188 -0.197593792
8193  3.385398105
8202 -0.941085395
8357  1.538241566
8367 -5.252052597
8477  0.275638125
8531 -0.087629889
8627 -1.765586364
8628  3.730107535
8707 -0.124221324
8775 -2.047327381
8800 -2.918034877
8854 -5.318113065
8857  1.251474449
8874  0.174438121
8946 -1.272022557
8983 -0.835762264
9021  0.475439588
9104  2.100936574
9158 -2.748330976
9198  4.205222430
9225  1.155010580
9292 -0.679532109
9340 -0.417469272
9347  0.330373558
9359  1.541003120
9397 -2.250668703
9508  1.020549481
9550 -1.337015597
9586  0.637601978

with conditional variances for "id"
```

```
ranova(model3)
```

```
ANOVA-like table for random-effects: Single term deletions

Model:
mathach ~ ses_grandmc + (1 | id)
         npar logLik   AIC    LRT Df Pr(>Chisq)
<none>      4 -23323 46653
(1 | id)    3 -23552 47110 458.92  1  < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
## Use BruceR package
HLM_summary(model3,test.rand = T, digits = 3)
```

```
Hierarchical Linear Model (HLM)
(also known as) Linear Mixed Model (LMM)
(also known as) Multilevel Linear Model (MLM)

Model Information:
Formula: mathach ~ ses_grandmc + (1 | id)
Level-1 Observations: N = 7185
Level-2 Groups/Clusters: id, 160

Model Fit:
AIC = 46653.169
BIC = 46680.688
R_(m)² = 0.07665  (Marginal R²: fixed effects)
R_(c)² = 0.18197  (Conditional R²: fixed + random effects)
Omega² = 0.23192  (= 1 - proportion of unexplained variance)

ANOVA Table:
─────────────────────────────────────────────────────────────
               Sum Sq  Mean Sq NumDF   DenDF      F     p
─────────────────────────────────────────────────────────────
ses_grandmc  18930.65 18930.65  1.00 6838.08 511.16 <.001 ***
─────────────────────────────────────────────────────────────

Fixed Effects:
Unstandardized Coefficients (b or γ):
Outcome Variable: mathach
───────────────────────────────────────────────────────────────────
                b/γ    S.E.     t     df     p      [95% CI of b/γ]
───────────────────────────────────────────────────────────────────
(Intercept)  12.658 (0.188) 67.33  148.3 <.001 *** [12.286, 13.029]
ses_grandmc   2.390 (0.106) 22.61 6838.1 <.001 *** [ 2.183,  2.597]
───────────────────────────────────────────────────────────────────
'df' is estimated by Satterthwaite approximation.

Standardized Coefficients (β):
Outcome Variable: mathach
────────────────────────────────────────────────────────────────
                 β    S.E.     t     df     p      [95% CI of β]
────────────────────────────────────────────────────────────────
ses_grandmc  0.271 (0.012) 22.61 6838.1 <.001 *** [0.247, 0.294]
────────────────────────────────────────────────────────────────

Random Effects:
──────────────────────────────────────────
 Cluster  K   Parameter   Variance     ICC
──────────────────────────────────────────
 id       160 (Intercept)  4.76817 0.11406
 Residual                 37.03440
──────────────────────────────────────────

ANOVA-like table for random-effects: Single term deletions

Model:
mathach ~ ses_grandmc + (1 | id)
         npar logLik   AIC    LRT Df Pr(>Chisq)
<none>      4 -23323 46653
(1 | id)    3 -23552 47110 458.92  1  < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# lrtest
anova(model1, model3)
```

```
refitting model(s) with ML (instead of REML)
```

```
Data: data_merged
Models:
model1: mathach ~ 1 + (1 | id)
model3: mathach ~ ses_grandmc + (1 | id)
       npar   AIC   BIC logLik deviance  Chisq Df Pr(>Chisq)
model1    3 47122 47142 -23558    47116
model3    4 46649 46677 -23321    46641 474.81  1  < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Model 3. One-Way ANCOVA (Cont'd)

```
Hierarchical Linear Model (HLM)
(also known as) Linear Mixed Model (LMM)
(also known as) Multilevel Linear Model (MLM)

Model Information:
Formula: mathach ~ ses_grandmc + (1 ¦ id)
Level-1 Observations: N = 7185
Level-2 Groups/Clusters: id, 160

Model Fit:
AIC = 46653.169
BIC = 46680.688
R_(m)² = 0.07665  (Marginal R²: fixed effects)
R_(c)² = 0.18197  (Conditional R²: fixed + random effects)
Omega² = 0.23192  (= 1 - proportion of unexplained variance)

ANOVA Table:
────────────────────────────────────────────────────────────
              Sum Sq  Mean Sq NumDF   DenDF      F     p
────────────────────────────────────────────────────────────
ses_grandmc  18930.65 18930.65  1.00 6838.08 511.16 <.001 ***
────────────────────────────────────────────────────────────

Fixed Effects:
Unstandardized Coefficients (b or γ):
Outcome Variable: mathach
──────────────────────────────────────────────────────────────────
                b/γ    S.E.      t     df     p       [95% CI of b/γ]
──────────────────────────────────────────────────────────────────
(Intercept)  12.658 (0.188) 67.33  148.3 <.001 *** [12.286, 13.029]
ses_grandmc   2.390 (0.106) 22.61 6838.1 <.001 *** [ 2.183,  2.597]
──────────────────────────────────────────────────────────────────
'df' is estimated by Satterthwaite approximation.
```

```
Standardized Coefficients (β):
Outcome Variable: mathach
──────────────────────────────────────────────────────────────
                 β    S.E.      t     df     p       [95% CI of β]
──────────────────────────────────────────────────────────────
ses_grandmc  0.271 (0.012) 22.61 6838.1 <.001 *** [0.247, 0.294]
──────────────────────────────────────────────────────────────

Random Effects:
──────────────────────────────────────────────
 Cluster  K   Parameter   Variance     ICC
──────────────────────────────────────────────
 id       160 (Intercept)  4.76817 0.11406
 Residual                 37.03440
──────────────────────────────────────────────

ANOVA-like table for random-effects: Single term deletions

Model:
mathach ~ ses_grandmc + (1 ¦ id)
           npar logLik   AIC    LRT Df Pr(>Chisq)
<none>        4 -23323 46653
(1 ¦ id)      3 -23552 47110 458.92  1  < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Model 3. One-Way ANCOVA (Cont'd)

- 두개의 고정효과 값 추정: γ_{00} = 12.658, γ_{10} = 2.390
- SES = 0인 사람들에 대한 수학 성취도 값이 γ_{00}=12.658으로 추정되는데 이는 평균적인 SES를 가진 사람들의 수학 성적임
- Level 1 독립변수에 의해서 설명되는 Level 1 분산의 비율

$$R^2_{Level-1} = \frac{\sigma^2_{\epsilon_1} - \sigma^2_{\epsilon_2}}{\sigma^2_{\epsilon_1}} = \frac{39.148 - 37.034}{39.148} = 0.054$$

- Level 1 독립변수에 의해서 설명되는 전체 분산의 비율

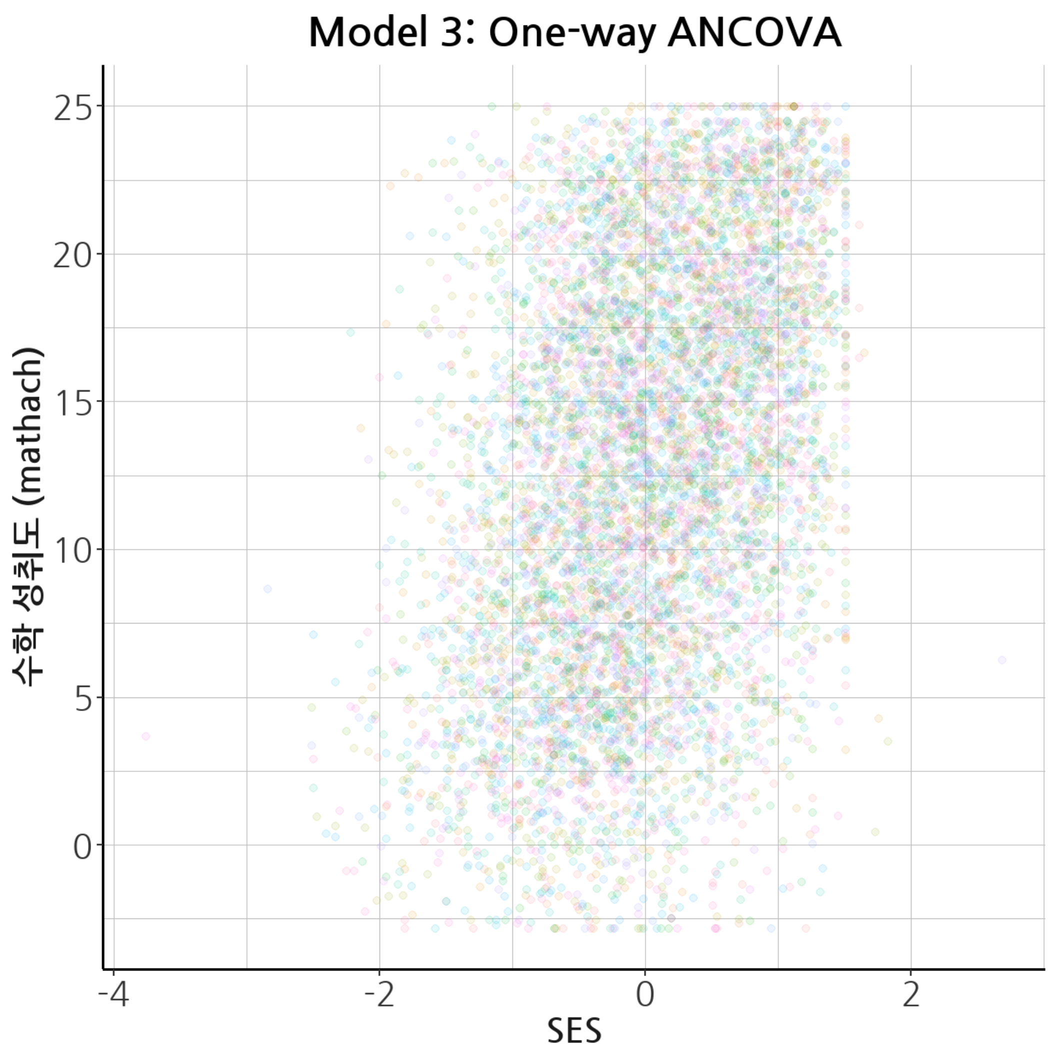
$$R^2_{tot} = \frac{(\sigma^2_{v_01} + \sigma^2_{\epsilon_1}) - (\sigma^2_{v_02} + \sigma^2_{\epsilon_2})}{\sigma^2_{v_01} + \sigma^2_{\epsilon_1}} = \frac{(8.553 + 39.148) - (4.768 + 37.034)}{(8.553 + 39.148)} = 0.124$$

- Students' SES explained 12% of the total variance in math achievement. (ICC보다 설명된 분산에 해당하는 값을 제시하는 것이 HLM을 정당화하기에 더 적절할 수 있음)
- $\bar{X}_{.j}$: Group Mean, $\bar{X}_{..}$: Grand Mean

$$\beta_{0j} = u_j - \beta_{1j}(\bar{X}_{.j} - \bar{X}_{..})$$

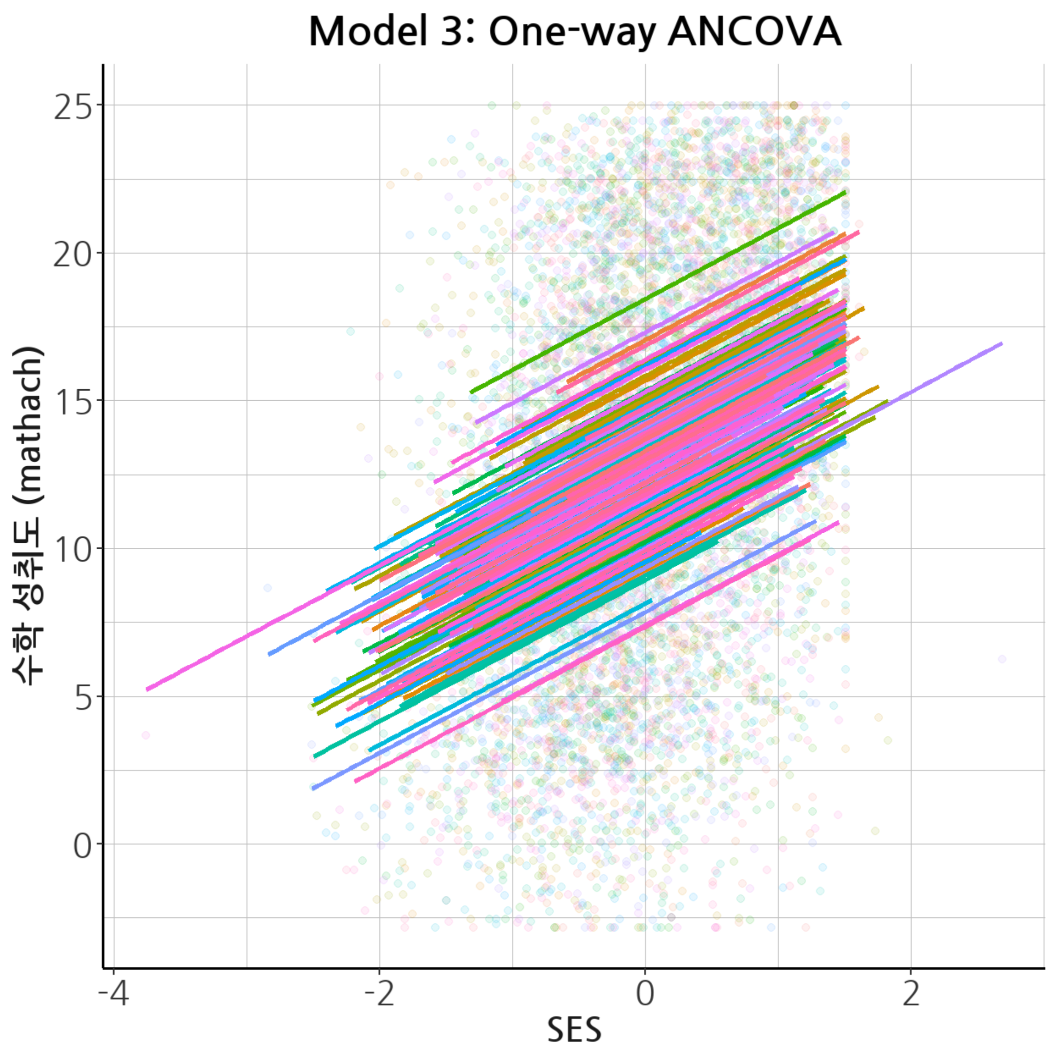
- Grand-mean centered 모형에서 절편은 각 집단의 평균에서 predictor의 집단평균과 전체평균의 편차를 반영한 adjustment 를 뺀 값

Model 3. One-Way ANCOVA: Visulization



Model 3. One-Way ANCOVA: Visulization

```
`geom_smooth()` using formula = 'y ~ x'
```



Model 4. Random-Coefficient Model

- Level-1의 절편과 기울기 모두를 random하게 변화하도록 모델링하지만 그러나 절편과 기울기의 random성에서 기인하는 구체적 variation에 대한 예측이나 모델링은 하지 않음 (hence “unconditional” model (NOT fully !))
- 목적: To investigate whether effects of level-1 predictors vary between level-2 units

e.g. SES에 의한 학생들간 수학 성취도 수준 차이가 학교별로 서로 다르게 나타나는가?

- level-1 predictors should be group-mean centered
  - Level 1: $Y_{ij} = \beta_{0j} + \beta_{1j}X_{ij} + \epsilon_{ij} \quad \epsilon_{ij} \sim N(0, \sigma^2)$
  - Level 2 intercept: $\beta_{0i} = \gamma_{00} + u_{0j} \quad u_{0j} \sim N(0, \tau_{00})$
  - **Level 2 slope**: $\beta_{1j} = \gamma_{10} + u_{1j} \quad u_{1j} \sim N(0, \tau_{11})$
- β_{0j}: 집단 j의 절편, β_{1j}: 집단 j의 기울기
- γ_{00}: Level-2 집단들의 평균 절편, γ_{10}: Level-2 집단들의 평균 기울기
- **해석**: τ_{00} 값이 크다면 학교 간 수학성취도 평균에 격차가 큼, τ_{11}의 값이 크다면 학교에 따라 가정배경(SES)으로 학생들의 수학성취도가 차별되는 수준에 큰 차이가 있다는 것, $\tau_{01} > 0$이러면 학교 평균(β_{0j})이 높은 학교에서 SES 에 의한 차별효과(β_{1j})가 더 큰 것이고 음수면 그 반대이다.

Model 4. Random-Coefficient Model (Cont'd)

$$Var = \begin{bmatrix} u_{0j} \\ u_{1j} \end{bmatrix} = \begin{bmatrix} \tau_{00} & \tau_{01} \\ \tau_{10} & \tau_{11} \end{bmatrix} = \mathrm{T}$$

- $Var(u_{0j})$ = τ_{00} = unconditional variance in level-1 intercepts
- $Var(u_{1j})$ = τ_{11} = unconditional variance in level-1 slopes
- $Cov(u_{0j}, u_{1j})$ = τ_{01} = ρ_{01} = unconditional variance between level-1 intercepts and slopes (association between mean school achievement and SES effect on achievement)

```
data_merged <- data_merged %>%
  group_by(id) %>%
  mutate(meanses = mean(ses),
         ses_groupmc = ses - meanses)

model4 <- lmer(mathach ~  ses_groupmc + (ses_groupmc | id), data=data_merged)
model4_alt <- lmer(mathach ~  ses_groupmc + (1 | id), data=data_merged)
icc(model4)
```

```
# Intraclass Correlation Coefficient

    Adjusted ICC: 0.197
  Unadjusted ICC: 0.188
```

```
# the individual random effects (level 2 residuals of the intercept, i.e. the υ0i)
ranef(model4)
```

```
$id
     (Intercept)   ses_groupmc
1224 -2.67923989  0.0696411630
1288  0.74984322  0.1827677450
1296 -4.59710183 -0.3209497448
1308  2.98588001 -0.1430037614
1317  0.49657768 -0.1960677035
1358 -1.24874052  0.5451934717
1374 -2.52333497  0.3206785928
1433  6.31910717 -0.0349847138
1436  4.99467763 -0.1031099856
1461  3.73573395  0.9315495349
1462 -1.99649948 -0.7996174417
1477  1.48919190 -0.3265526778
1499 -4.60577634  0.4717593269
1637 -4.85039299  0.1841505688
1906  3.09963618 -0.0006880699
1909  1.55388144  0.1427354906
1942  4.77581165 -0.2461761138
1946  0.24799828  0.3622359799
2030 -0.51319378 -0.2259320128
2208  2.58678959  0.1364348461
2277 -3.12831382 -1.4152304129
2305 -1.41364145 -1.0436412618
2336  3.56073343 -0.0615371692
2458  1.25742050  0.2448031006
2467 -2.30057875  0.1803411279
2526  4.10912742 -0.5596790741
2626  0.68659534  0.3460826555
2629  2.11154164 -0.6736704306
2639 -5.47405026 -0.6686461886
2651 -1.39242986  0.5968870322
2655 -0.26491007  0.8324840851
2658  0.69533535  0.1138389549
2755  3.52103511 -0.4042355387
2768 -1.49277354  0.3482872530
2771 -0.73342173  0.4372051380
2818  1.12440317  0.1449145976
2917 -4.24281298 -0.4320950519
2990  5.33989637 -0.2259789819
2995 -2.83169173 -0.3143349119
3013 -0.02197447  0.3038862361
3020  1.64067177 -0.1659524975
3039  3.60376804  0.1371175984
3088 -3.14979101 -0.1378284422
3152  0.53096122  0.2368664728
3332  1.47736055 -0.0244529244
3351 -1.05589543  0.0843532167
3377 -3.15806886 -0.8501479292
3427  6.51296025 -0.7405067393
3498  3.47404222 -0.5585393513
3499  0.57446103 -0.2583467872
3533 -2.04965225 -0.5203207403
3610  2.55130427  0.2549678261
3657 -2.87363385  0.5522268457
3688  1.83847079 -0.1166882464
3705 -2.10826479 -0.2996044265
3716 -2.04786455  1.2125054018
3838  3.17535913 -0.4877983317
3881 -0.62253409  0.0323222784
3967 -0.55435299  0.3059336842
```

```
3992  1.85840340 -0.4310641254
3999 -1.54646330  0.5657920799
4042  1.57449130 -0.1558325365
4173  0.08241410  0.2884781255
4223  1.81622077  0.0761304100
4253 -3.00839230 -0.8997423933
4292  0.21093915 -0.7972607480
4325  0.56030116  0.2237309938
4350 -0.68851892  0.4643070946
4383 -0.99597696  0.5304788682
4410  0.75935484  0.1293909980
4420  1.09478710  0.1697220034
4458 -6.27387400 -0.3118438618
4511  0.71770099 -0.5712561222
4523 -3.93050953  0.0433590047
4530 -3.35603097 -0.1823467576
4642  1.83718805  0.3682753927
4868 -0.29157715 -0.2181151263
4931  1.07420402 -0.4243587904
5192 -1.93585263 -0.1525786098
5404  2.58594121 -0.1963777818
5619  2.61646758  0.9436473544
5640  0.48974832  0.4344636095
5650  1.49377915 -0.5009162141
5667  1.06981595  0.4172415022
5720  1.52490374  0.0883371177
5761 -1.38395795  0.2954439475
5762 -7.46240892 -0.5251043518
5783  0.45060465  0.2202889757
5815 -4.58776858  0.0848651423
5819 -0.45881149 -0.0572221592
5838  0.92649828 -0.0570348658
5937  3.61157842 -0.1532109472
6074  1.06175146 -0.1886748702
6089  2.59935202 -0.1050006397
6144 -3.72385472  0.1534854037
6170  1.29196807  0.5377376533
6291 -2.25378819  0.3385653388
6366  2.81400478 -0.2057118666
6397  0.15036375  0.2220575201
6415 -0.71754928  0.4248189203
6443 -2.77624753 -0.7324371126
6464 -4.84084672 -0.2268469810
6469  5.41706332 -0.1078840229
6484  0.24369898 -0.3757685158
6578 -0.59684051  0.0547168430
6600 -0.86282764  0.8467171434
6808 -3.05621042  0.0176237798
6816  1.76489678 -0.2771872946
6897  2.26840386  0.4723836927
6990 -6.16997384 -0.5408173094
7011  1.04784024  0.5491629979
7101 -0.68512992 -0.2306366575
7172 -4.17070947 -0.3413563793
7232 -0.08308717  0.6772868543
7276  0.04300508  0.6010380105
7332  1.83842943  0.0837804514
7341 -2.62546194 -0.2077004984
7342 -1.37130057 -0.3075617975
7345 -1.20257562  0.8336154981
7364  1.39932956 -0.3262436472
7635  2.24375103  0.0853201373
```

```
7688  5.36352489 -0.4805869044
7697  2.72683185  0.1823900683
7734 -1.73181006  0.9130040230
7890 -3.96995314 -0.7271962982
7919  1.98879506  0.3033424421
8009  1.32821466 -0.1319051654
8150  2.01860642 -0.5464869295
8165  3.51143771 -0.0900062582
8175 -0.83256585 -0.1345415330
8188  0.09554245  0.4458615795
8193  3.27431431  0.0435391144
8202 -0.82154447  0.3621316642
8357  1.51080894  0.1701178475
8367 -6.21225509 -0.2682931396
8477 -0.09884061  0.4946028908
8531  0.81090711  0.2965669678
8627 -1.62353142 -0.1126353166
8628  3.63873380 -0.2513930420
8707  0.23007562  0.4381167495
8775 -2.91448780 -0.3506936250
8800 -4.68093514  0.0765319704
8854 -7.41704601 -0.0988333481
8857  2.49439510 -0.3407890965
8874 -0.51668392  0.4776347189
8946 -2.10834376 -0.1973615913
8983 -1.51947790 -0.2302489564
9021  1.91623425  0.0946454438
9104  3.89582186 -0.1021599188
9158 -3.78582512  0.5773986262
9198  5.68191577  0.1055898242
9225  1.81895327  0.1890745771
9292 -1.92911873 -0.1302723646
9340 -1.27075942  0.1589688690
9347  0.84099766  0.1668696787
9359  2.43611992 -0.7283666986
9397 -2.09214935  0.0526511572
9508  0.83953216  0.3082689211
9550 -1.34662145  0.4120808038
9586  2.07789207 -0.1376401945

with conditional variances for "id"
```

```
ranova(model4)
```

```
ANOVA-like table for random-effects: Single term deletions

Model:
mathach ~ ses_groupmc + (ses_groupmc | id)
                                  npar logLik   AIC    LRT Df Pr(>Chisq)
<none>                               6 -23357 46726
ses_groupmc in (ses_groupmc | id)    4 -23362 46732 9.7617  2   0.007591 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
## Use BruceR package
HLM_summary(model4,test.rand = T, digits = 3)
```

```
Hierarchical Linear Model (HLM)
(also known as) Linear Mixed Model (LMM)
(also known as) Multilevel Linear Model (MLM)

Model Information:
Formula: mathach ~ ses_groupmc + (ses_groupmc | id)
Level-1 Observations: N = 7185
Level-2 Groups/Clusters: id, 160

Model Fit:
AIC = 46726.234
BIC = 46767.513
R_(m)² = 0.04393  (Marginal R²: fixed effects)
R_(c)² = 0.23194  (Conditional R²: fixed + random effects)
Omega² = 0.24448  (= 1 - proportion of unexplained variance)

ANOVA Table:
────────────────────────────────────────────────────────────
               Sum Sq  Mean Sq NumDF  DenDF      F     p
────────────────────────────────────────────────────────────
ses_groupmc  10731.21 10731.21  1.00 155.22 292.40 <.001 ***
────────────────────────────────────────────────────────────


Fixed Effects:
Unstandardized Coefficients (b or γ):
Outcome Variable: mathach
──────────────────────────────────────────────────────────────────
                b/γ    S.E.     t    df     p      [95% CI of b/γ]
──────────────────────────────────────────────────────────────────
(Intercept)  12.636 (0.245) 51.68 156.8 <.001 *** [12.153, 13.119]
ses_groupmc   2.193 (0.128) 17.10 155.2 <.001 *** [ 1.940,  2.447]
──────────────────────────────────────────────────────────────────
'df' is estimated by Satterthwaite approximation.

Standardized Coefficients (β):
Outcome Variable: mathach
───────────────────────────────────────────────────────────────
                 β    S.E.     t    df     p      [95% CI of β]
───────────────────────────────────────────────────────────────
ses_groupmc  0.211 (0.012) 17.10 155.2 <.001 *** [0.186, 0.235]
───────────────────────────────────────────────────────────────


Random Effects:
──────────────────────────────────────────
 Cluster  K   Parameter   Variance     ICC
──────────────────────────────────────────
 id       160 (Intercept)  8.68104 0.19129
              ses_groupmc  0.69400
 Residual                 36.70019
──────────────────────────────────────────


ANOVA-like table for random-effects: Single term deletions

Model:
mathach ~ ses_groupmc + (ses_groupmc | id)
                                  npar logLik   AIC    LRT Df Pr(>Chisq)
<none>                               6 -23357 46726
ses_groupmc in (ses_groupmc | id)    4 -23362 46732 9.7617  2   0.007591 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# lrtest for coefficient randomness
anova(model4, model4_alt)
```

```
refitting model(s) with ML (instead of REML)
```

```
Data: data_merged
Models:
model4_alt: mathach ~ ses_groupmc + (1 | id)
model4: mathach ~ ses_groupmc + (ses_groupmc | id)
           npar   AIC   BIC logLik deviance Chisq Df Pr(>Chisq)
model4_alt    4 46728 46756 -23360    46720
model4        6 46723 46764 -23356    46711 9.433  2   0.008946 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Model 4. Random-Coefficient Model (Cont'd)

```
Model Information:
Formula: mathach ~ ses_groupmc + (ses_groupmc ¦ id)
Level-1 Observations: N = 7185
Level-2 Groups/Clusters: id, 160

Model Fit:
AIC = 46726.234
BIC = 46767.513
R_(m)² = 0.04393  (Marginal R²: fixed effects)
R_(c)² = 0.23194  (Conditional R²: fixed + random effects)
Omega² = 0.24448  (= 1 - proportion of unexplained variance)

ANOVA Table:
───────────────────────────────────────────────────────────
               Sum Sq  Mean Sq NumDF  DenDF      F     p
───────────────────────────────────────────────────────────
ses_groupmc  10731.21 10731.21  1.00 155.22 292.40 <.001 ***
───────────────────────────────────────────────────────────

Fixed Effects:
Unstandardized Coefficients (b or γ):
Outcome Variable: mathach
─────────────────────────────────────────────────────────────────
                b/γ    S.E.     t    df     p      [95% CI of b/γ]
─────────────────────────────────────────────────────────────────
(Intercept)  12.636 (0.245) 51.68 156.8 <.001 *** [12.153, 13.119]
ses_groupmc   2.193 (0.128) 17.10 155.2 <.001 *** [ 1.940,  2.447]
─────────────────────────────────────────────────────────────────
'df' is estimated by Satterthwaite approximation.

Standardized Coefficients (β):
Outcome Variable: mathach
──────────────────────────────────────────────────────────────
                 β    S.E.     t    df     p      [95% CI of β]
──────────────────────────────────────────────────────────────
ses_groupmc  0.211 (0.012) 17.10 155.2 <.001 *** [0.186, 0.235]
──────────────────────────────────────────────────────────────

Random Effects:
──────────────────────────────────────────
 Cluster  K   Parameter   Variance     ICC
──────────────────────────────────────────
 id       160 (Intercept)  8.68104 0.19129
              ses_groupmc  0.69400
 Residual                 36.70019
──────────────────────────────────────────

ANOVA-like table for random-effects: Single term deletions

Model:
mathach ~ ses_groupmc + (ses_groupmc ¦ id)
                                  npar logLik   AIC    LRT Df Pr(>Chisq)
<none>                               6 -23357 46726
ses_groupmc in (ses_groupmc ¦ id)    4 -23362 46732 9.7617  2   0.007591 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Model 4. Random-Coefficient Model (Cont'd)

1. Fixed effects:
   - $\hat{\gamma_{00}}$ = 12.64 (the average school mean)
   - $\hat{\gamma_{10}}$ = 2.19 (the average SES-achievement)
2. Random effects
   - Level 1: (σ^2) = 36.70
   - Level 2 intercept: $(Var(u_{0j}) = \tau_{00})$ = 8.681
   - Level 2 slope: $(Var(u_{1j}) = \tau_{11})$ = 0.69
   - $Var(u_{0j})$: The estimated variance among the means

- $Var(u_{1j})$: The estimated variance of slopes
- Range of plausible values:

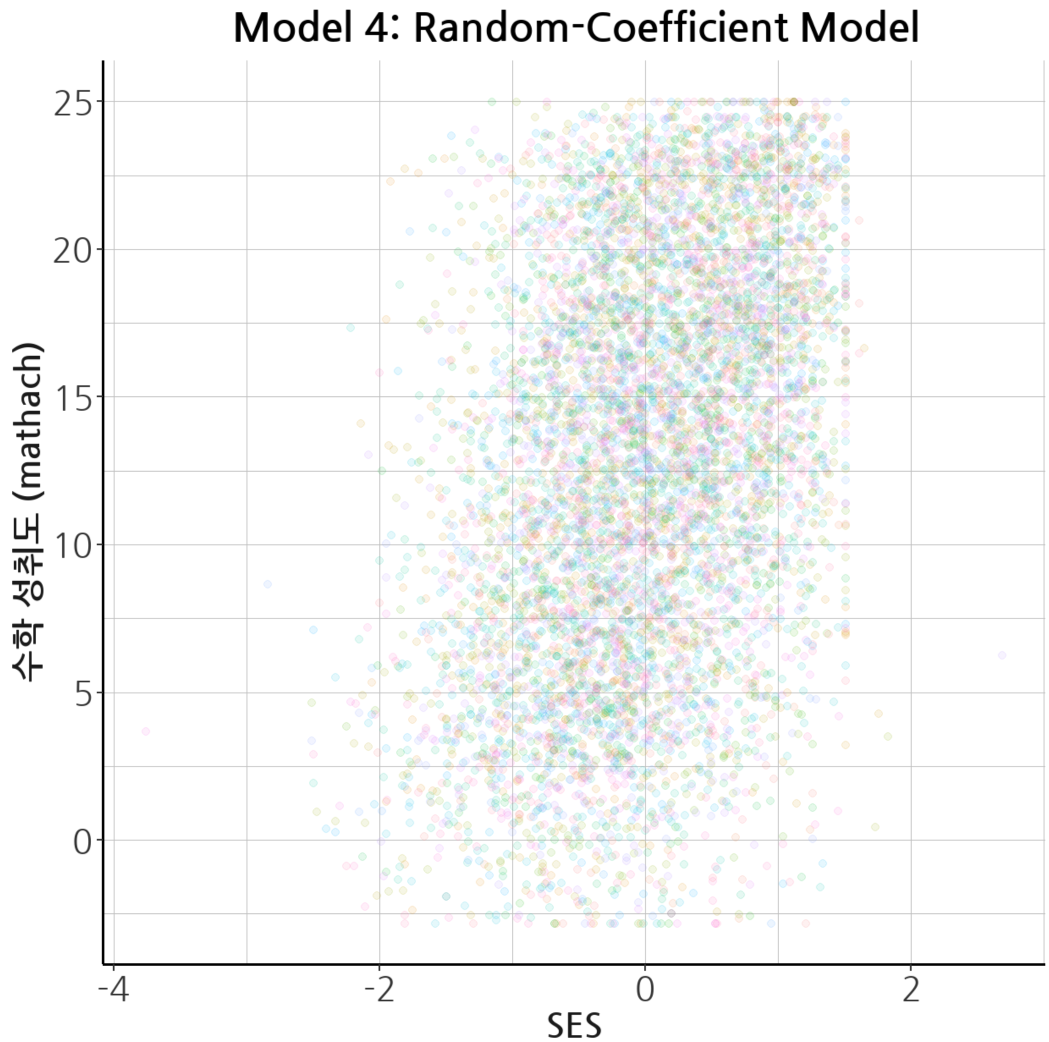
$$Intercept : 12.64 \pm 1.96\sqrt{8.68}, \quad Slope : 2.193 \pm 1.96\sqrt{0.69}$$

3. Variance Explained (at level 1)

$$\frac{\hat{\sigma}(Random\ ANOVA) - \hat{\sigma}(SES)}{\hat{\sigma}(Random\ ANOVA)} = \frac{39.148 - 36.70019}{39.148} = 0.063$$

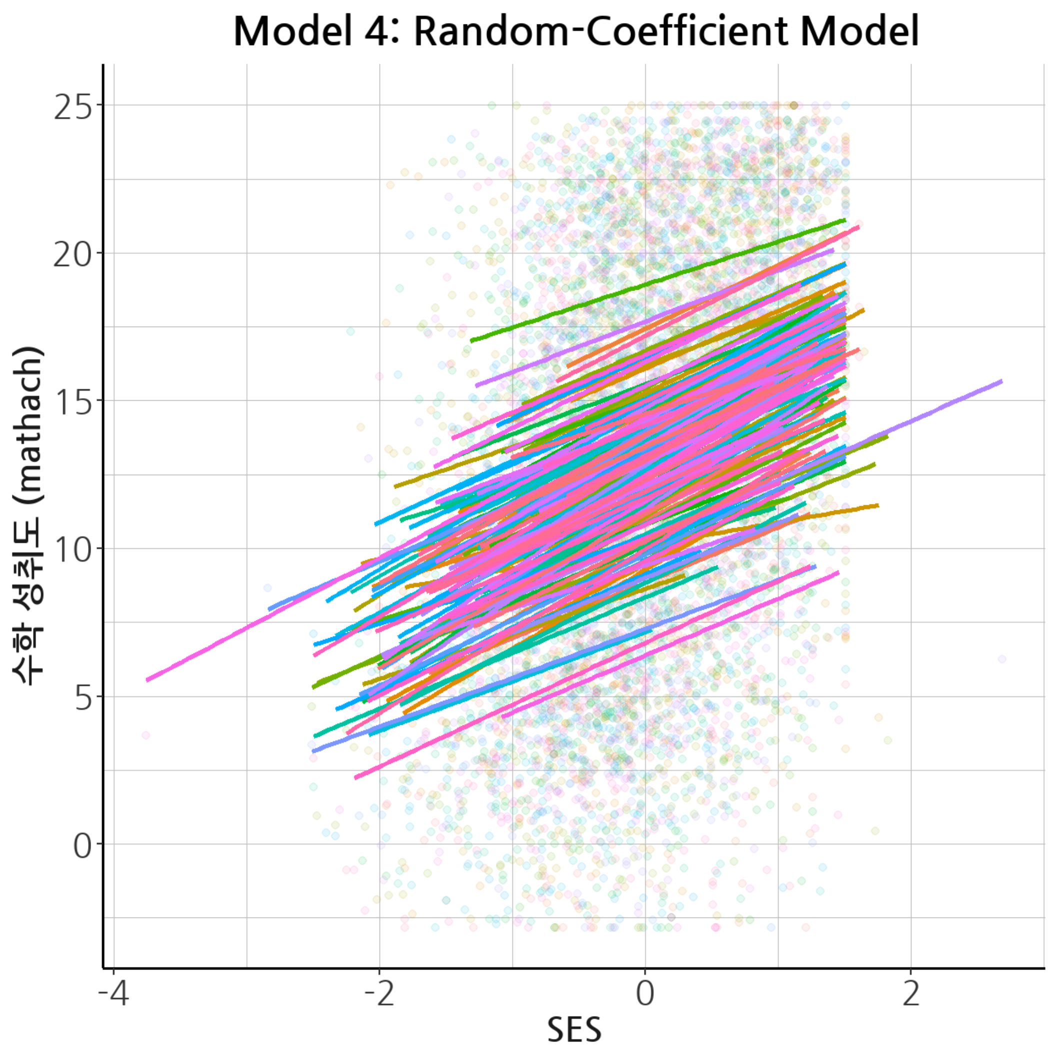
SES explains 6.3% of student-level variance

Model 4. Random-Coefficient Model: Visulization



Model 4. Random-Coefficient Model: Visulization

```
`geom_smooth()` using formula = 'y ~ x'
```



Model 5-1. Intercepts-as-Outcomes Model

- Intercepts-as-Outcomes의 경우 Random-effects ANCOVA with Level-2 predictor 혹은 Contextual Effect Model로 볼 수 있음

1. Random-effects ANCOVA with Level-2 predictor
    - 기본적으로는 Model 3와 동일하며, Level-2 predictor가 추가적으로 투입되어 추정이 이루어짐
    - level-1 covariate 는 fixed effects (i.e., assuming that the effects of these covariates are the same for all schools) 일수도 random effects (i.e., assuming that the effects of level 1 variables vary across schools) 일수도 있음
    - i.e. **with** a random intercept, and **with or without** a random slope
    - Level 1: $Y_{ij} = \beta_{0j} + \beta_{1j}X_{ij} + \epsilon_{ij} \quad \epsilon_{ij} \sim N(0, \sigma^2)$
    - Level 2: $\beta_{0i} = \gamma_{00} + \gamma_{01}Z_j + u_{0j}$, $\beta_{1j} = \gamma_{10} \quad u_{0j} \sim N(0, \tau_{00})$
    - Fixed Effect: γ_{00} = average outcome for sample of groups, γ_{01} = average group effect (intercept) on outcome, γ_{10} = average individual effect (slope) on outcome
    - Random Effect: ϵ_{ij} = residuals, u_{0j} = unique effect of group j on average outcome

Model 5-1. Intercepts-as-Outcomes Model (Cont'd)

1. **Random coefficient not included**
    - Level 1: $Y_{ij} = \beta_{0j} + \beta_{1j}SES_{ij} + \epsilon_{ij} \quad \epsilon_{ij} \sim N(0, \sigma^2)$
    - Level 2 intercept: $\beta_{0i} = \gamma_{00} + \gamma_{01}SECTOR_j + u_{0j} \quad u_{0j} \sim N(0, \tau_{00})$

- Level 2 slope: $\beta_{1j} = \gamma_{10}$

2. **Random coefficient included**
   - Level 1: $Y_{ij} = \beta_{0j} + \beta_{1j}SES_{ij} + \epsilon_{ij} \quad \epsilon_{ij} \sim N(0, \sigma^2)$
   - Level 2 intercept: $\beta_{0i} = \gamma_{00} + \gamma_{01}SECTOR_j + u_{0j} \quad u_{0j} \sim N(0, \tau_{00})$
   - Level 2 slope: $\beta_{1j} = \gamma_{10} + u_{1j} \quad u_{1j} \sim N(0, \tau_{11})$

Model 5-1. Intercepts-as-Outcomes Model (Cont'd)

- Random coefficient not included

```
## Random-effects ANCOVA with Level-2 predictor

# without random coefficient
model5 <- lmer(mathach ~  ses + sector + (1 | id), data=data_merged)
icc(model5)
```

```
# Intraclass Correlation Coefficient

    Adjusted ICC: 0.090
  Unadjusted ICC: 0.080
```

```
# the individual random effects (level 2 residuals of the intercept, i.e. the υ0i)
ranef(model5)
```

```
$id
      (Intercept)
1224 -0.800702281
1288  1.072112010
1296 -2.540549636
1308  0.786611650
1317 -1.208978410
1358 -0.349037144
1374 -1.442598746
1433  3.269681163
1436  2.405623033
1461  2.694363392
1462 -1.474541875
1477  0.025613409
1499 -2.481529960
1637 -2.200523758
1906  0.797278230
1909  1.556580363
1942  3.544134664
1946  0.938141041
2030 -0.346446003
2208  0.496803980
2277 -2.329023750
2305 -1.035352069
2336  3.089512369
2458 -0.318784175
2467 -0.660676805
2526  2.088649117
2626  1.448551619
2629  1.202816150
2639 -2.269984127
2651 -0.620268902
2655  1.921165752
2658 -1.197344423
2755  1.100905135
2768 -0.502632559
2771  0.787416627
2818  1.529793534
2917 -1.494958019
2990  2.541309847
2995 -1.195214586
3013  0.838028535
3020  0.024797147
3039  1.567963455
3088 -1.153476327
3152  1.186998955
3332  0.875591607
3351 -0.817461314
3377 -0.933667179
3427  4.590823462
3498  0.844355699
3499 -1.274496642
3533 -2.556171709
3610  1.080296055
3657 -0.548108359
3688 -0.101562924
3705 -3.306764977
3716 -0.282096431
3838  1.600126345
3881 -0.058406863
3967  0.712952956
```

```
3992 -0.035512822
3999 -0.458486756
4042 -0.396663900
4173 -0.953316477
4223  0.838765152
4253 -2.999674238
4292  0.172423639
4325  1.376697751
4350 -0.173145056
4383 -0.304599933
4410  1.224900525
4420  2.045365629
4458 -2.824033223
4511 -0.133205430
4523 -4.399557530
4530 -2.886166623
4642  2.237347466
4868 -1.827034917
4931 -0.762925659
5192 -2.398130914
5404 -0.309316295
5619  0.519254680
5640  1.581672366
5650  0.327372362
5667 -0.889996114
5720  0.323816071
5761 -1.604529654
5762 -3.584955088
5783  0.757872192
5815 -2.020466754
5819 -0.009371509
5838  1.207527881
5937  2.607510222
6074  0.522729384
6089  2.795253118
6144 -1.730370418
6170  2.149904585
6291 -0.289530833
6366  0.938445155
6397  1.363000701
6415  0.494032165
6443 -1.052494193
6464 -2.243715942
6469  2.421897425
6484  1.268803097
6578 -0.507883698
6600  0.095056678
6808 -1.794487683
6816 -0.454465908
6897  2.114854651
6990 -3.849270222
7011  0.052307981
7101  0.084181041
7172 -4.122921302
7232  0.869639989
7276  0.651228192
7332  0.091840387
7341 -1.282373824
7342 -1.355056334
7345 -0.389421737
7364  0.459617666
7635  0.558677752
```

```
7688  3.508185831
7697  2.579447183
7734 -0.030660909
7890 -1.784801696
7919  1.607546556
8009 -0.790433807
8150  0.249329036
8165  1.574821064
8175  0.296474514
8188  0.518324687
8193  2.295384433
8202 -0.163169900
8357  2.126190264
8367 -4.206788796
8477  0.997953377
8531  0.675155710
8627 -0.911315552
8628  2.610507166
8707  0.658723931
8775 -1.195065349
8800 -3.652392985
8854 -4.323980546
8857  0.215287508
8874  0.895390450
8946 -0.434023014
8983 -0.024244160
9021 -0.518179102
9104  1.054227094
9158 -1.865065295
9198  3.099340487
9225  1.834529391
9292 -0.018278673
9340  0.299383246
9347 -0.666137675
9359  0.512805813
9397 -1.385644030
9508  0.064190164
9550 -0.561210889
9586 -0.368357789

with conditional variances for "id"
```

```
ranova(model5)
```

```
ANOVA-like table for random-effects: Single term deletions

Model:
mathach ~ ses + sector + (1 | id)
         npar logLik   AIC   LRT Df Pr(>Chisq)
<none>      5 -23306 46621
(1 | id)    4 -23473 46955 335.6  1  < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
## Use BruceR package
HLM_summary(model5,test.rand = T, digits = 3)
```

```
Hierarchical Linear Model (HLM)
(also known as) Linear Mixed Model (LMM)
(also known as) Multilevel Linear Model (MLM)

Model Information:
Formula: mathach ~ ses + sector + (1 | id)
Level-1 Observations: N = 7185
Level-2 Groups/Clusters: id, 160

Model Fit:
AIC = 46621.151
BIC = 46655.550
R_(m)² = 0.11450  (Marginal R²: fixed effects)
R_(c)² = 0.19463  (Conditional R²: fixed + random effects)
Omega² = 0.23125  (= 1 - proportion of unexplained variance)

ANOVA Table:
────────────────────────────────────────────────────────
          Sum Sq  Mean Sq NumDF   DenDF      F     p
────────────────────────────────────────────────────────
ses     18768.30 18768.30  1.00 6738.86 506.75 <.001 ***
sector   1404.73  1404.73  1.00  147.36  37.93 <.001 ***
────────────────────────────────────────────────────────

Fixed Effects:
Unstandardized Coefficients (b or γ):
Outcome Variable: mathach
──────────────────────────────────────────────────────────────────────
                  b/γ    S.E.     t     df     p      [95% CI of b/γ]
──────────────────────────────────────────────────────────────────────
(Intercept)     11.719 (0.228) 51.39  153.6 <.001 *** [11.268, 12.169]
ses              2.375 (0.105) 22.51 6738.9 <.001 *** [ 2.168,  2.582]
sectorCatholic   2.101 (0.341)  6.16  147.4 <.001 *** [ 1.427,  2.775]
──────────────────────────────────────────────────────────────────────
'df' is estimated by Satterthwaite approximation.

Standardized Coefficients (β):
Outcome Variable: mathach
───────────────────────────────────────────────────────────────────
                   β    S.E.     t     df     p      [95% CI of β]
───────────────────────────────────────────────────────────────────
ses             0.269 (0.012) 22.51 6738.9 <.001 *** [0.246, 0.293]
sectorCatholic  0.153 (0.025)  6.16  147.4 <.001 *** [0.104, 0.202]
───────────────────────────────────────────────────────────────────

Random Effects:
──────────────────────────────────────────
 Cluster  K   Parameter   Variance     ICC
──────────────────────────────────────────
 id       160 (Intercept)  3.68504 0.09049
 Residual                 37.03691
──────────────────────────────────────────

ANOVA-like table for random-effects: Single term deletions

Model:
mathach ~ ses + sector + (1 | id)
         npar logLik   AIC   LRT Df Pr(>Chisq)
<none>      5 -23306 46621
(1 | id)    4 -23473 46955 335.6  1  < 2.2e-16 ***
```

```
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

- Random coefficient included

```
# with random coefficient

model6 <- lmer(mathach ~  ses + sector + (ses | id), data=data_merged)
icc(model6)
```

```
# Intraclass Correlation Coefficient

    Adjusted ICC: 0.103
  Unadjusted ICC: 0.090
```

```
# the individual random effects (level 2 residuals of the intercept, i.e. the u0i)
ranef(model6)
```

```
$id
     (Intercept)          ses
1224 -0.603522850 -0.0586182185
1288  1.287945952  0.2957946163
1296 -2.567509518 -0.5159846148
1308  0.609747717  0.0572618879
1317 -1.332047224 -0.3600839791
1358 -0.022268344  0.2481580769
1374 -1.189263751 -0.0498427152
1433  2.870458873  0.5645850185
1436  2.099424459  0.3453038474
1461  2.669990744  0.9398341826
1462 -2.165437642 -0.7089560115
1477 -0.175200637 -0.2380985979
1499 -2.161890742 -0.0037708804
1637 -1.971862355 -0.1162925349
1906  0.592557550  0.0816027095
1909  1.730382389  0.3582376089
1942  3.469645031  0.6264217596
1946  1.210074473  0.3501925342
2030 -0.157074470 -0.1657367033
2208  0.308177675  0.0911693949
2277 -3.494478527 -1.2079318610
2305 -1.825510471 -0.8456939656
2336  3.122561160  0.5217838468
2458 -0.461344021  0.0151192553
2467 -0.416561501  0.0208797636
2526  1.832573448  0.0132319133
2626  1.746718148  0.4329954277
2629  0.866284336 -0.2930210367
2639 -2.643796019 -0.6200768329
2651 -0.313193803  0.2241347990
2655  2.676034915  0.7231545462
2658 -1.292521990 -0.2124235185
2755  0.888822559 -0.0463754453
2768 -0.224175100  0.1223595304
2771  1.171738605  0.3700035983
2818  1.733366874  0.3394190630
2917 -1.665845627 -0.4506618255
2990  2.204820686  0.2911367302
2995 -1.157257448 -0.3641334787
3013  1.108785216  0.3116021556
3020 -0.153743637 -0.1452987401
3039  1.434765205  0.3164917840
3088 -1.074346497 -0.2319931295
3152  1.414082012  0.2888217630
3332  0.993599207  0.1738558247
3351 -0.601998638 -0.0903171688
3377 -1.222509675 -0.6447141387
3427  4.265030635  0.2888906342
3498  0.671108963 -0.1487718379
3499 -1.380455897 -0.4053773310
3533 -2.908826395 -0.7246447840
3610  0.918250466  0.2418883602
3657 -0.069841456  0.2911815165
3688 -0.257403143 -0.1232832548
3705 -3.424081301 -0.7586237970
3716  0.441239534  0.7140661890
3838  1.350751877 -0.0716296378
3881  0.137410302  0.0245457162
3967  1.019641597  0.2789992324
```

```
3992 -0.203232362 -0.2738949275
3999 -0.123707542  0.2664422986
4042 -0.534427315 -0.2066744335
4173 -1.066913465 -0.0511480338
4223  0.696860587  0.1192567234
4253 -3.623187167 -0.9868118923
4292 -0.363466648 -0.5424644819
4325  1.620567298  0.3069513910
4350  0.095145329  0.2225204788
4383  0.007443874  0.2384301153
4410  1.430730273  0.2820522671
4420  2.335184047  0.4042598221
4458 -2.961144078 -0.4130360313
4511 -0.439624202 -0.3919549595
4523 -4.591883049 -0.6828846142
4530 -3.299002457 -0.4868435597
4642  2.450023809  0.5413133115
4868 -1.943918185 -0.4831934907
4931 -0.894433275 -0.4051964454
5192 -2.632705750 -0.5056332825
5404 -0.391242724 -0.1922849649
5619  0.298133288  0.4863655843
5640  1.929033221  0.4799569331
5650  0.051538760 -0.2976795578
5667 -1.013708828 -0.0058924876
5720  0.164345416  0.0383145641
5761 -1.727576681 -0.0859689366
5762 -3.875490519 -0.5408755786
5783  0.968431335  0.2513221579
5815 -1.855701774 -0.1685615682
5819  0.179692835 -0.0203350855
5838  1.366463709  0.1968490339
5937  2.606548355  0.4682585654
6074  0.291729512 -0.0960259270
6089  2.945102102  0.4123307308
6144 -1.529073122 -0.1239558144
6170  2.615888977  0.6240257457
6291  0.057209050  0.1735415531
6366  0.734271993 -0.0152087252
6397  1.646660030  0.2943744089
6415  0.818978829  0.3131563030
6443 -1.235173952 -0.5759780039
6464 -2.275591004 -0.3833205213
6469  2.049999824  0.3247444774
6484  1.350758890 -0.0207854223
6578 -0.694978061 -0.0795998543
6600  0.468689023  0.4940192068
6808 -1.617710811 -0.2497065053
6816 -0.565253502 -0.2766171761
6897  2.241961431  0.5890945703
6990 -4.032795451 -0.7908456718
7011  0.038159654  0.2553720403
7101  0.185379025 -0.1110594731
7172 -4.532273765 -0.8168286414
7232  1.232975024  0.5120853229
7276  0.931096277  0.4279456475
7332 -0.067522765 -0.0008396883
7341 -1.168682787 -0.3116534287
7342 -1.721289797 -0.4065897197
7345 -0.049465957  0.4220253162
7364  0.219999857 -0.1437702397
7635  0.385807657  0.0749343270
```

```
7688  3.233149513  0.2886828516
7697  2.739172900  0.5669086438
7734  0.638530428  0.5553023127
7890 -1.992757419 -0.6564830670
7919  1.728202341  0.4604099661
8009 -0.894157384 -0.2575974047
8150  0.041634027 -0.2791971234
8165  1.355909303  0.1652849170
8175  0.437640449 -0.0244021585
8188  0.786216504  0.3320041566
8193  2.186475614  0.3231473020
8202  0.094334874  0.1744077551
8357  2.360406496  0.3873949350
8367 -4.265923482 -0.8618873017
8477  1.364550995  0.4261432318
8531  0.844078676  0.2725392682
8627 -0.721472697 -0.2028193292
8628  2.414264108  0.1845377219
8707  0.895461379  0.3363477086
8775 -1.166675260 -0.3678965968
8800 -3.888742440 -0.3710062094
8854 -4.334701712 -0.4896462176
8857  0.065931120 -0.1747805814
8874  1.300856419  0.4049851188
8946 -0.325698185 -0.1923720881
8983  0.095211601 -0.1375942657
9021 -0.635677547 -0.1084893193
9104  0.811426257  0.1090644336
9158 -1.491076759  0.1259596908
9198  2.821533707  0.5769478791
9225  2.000047761  0.4137947610
9292  0.068622334 -0.0612909023
9340  0.565105545  0.1474907744
9347 -0.805821811 -0.0797023817
9359  0.313597988 -0.3180153336
9397 -1.170207225 -0.1905277503
9508 -0.005389342  0.1332150231
9550 -0.281902122  0.1397871305
9586 -0.477811181 -0.1872385505

with conditional variances for "id"
```

```
ranova(model6)
```

```
ANOVA-like table for random-effects: Single term deletions

Model:
mathach ~ ses + sector + (ses | id)
                  npar logLik   AIC    LRT Df Pr(>Chisq)
<none>               7 -23301 46616
ses in (ses | id)    5 -23306 46621 9.2963  2   0.009579 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
## Use BruceR package
HLM_summary(model6, test.rand = T, digits = 3)
```

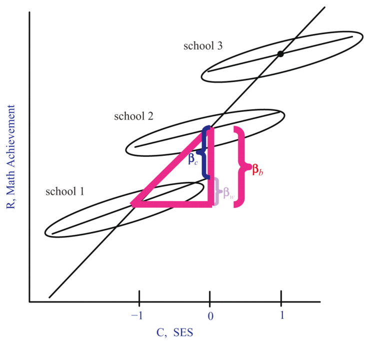
```
Hierarchical Linear Model (HLM)
(also known as) Linear Mixed Model (LMM)
(also known as) Multilevel Linear Model (MLM)

Model Information:
Formula: mathach ~ ses + sector + (ses | id)
Level-1 Observations: N = 7185
Level-2 Groups/Clusters: id, 160

Model Fit:
AIC = 46615.855
BIC = 46664.013
R_(m)² = 0.12693  (Marginal R²: fixed effects)
R_(c)² = 0.21691  (Conditional R²: fixed + random effects)
Omega² = 0.23873  (= 1 - proportion of unexplained variance)

ANOVA Table:
───────────────────────────────────────────────────────
           Sum Sq  Mean Sq NumDF  DenDF      F     p
───────────────────────────────────────────────────────
ses      15072.41 15072.41  1.00 157.84 409.57 <.001 ***
sector    2001.57  2001.57  1.00 151.31  54.39 <.001 ***
───────────────────────────────────────────────────────

Fixed Effects:
Unstandardized Coefficients (b or γ):
Outcome Variable: mathach
─────────────────────────────────────────────────────────────────────
                   b/γ    S.E.     t    df     p      [95% CI of b/γ]
─────────────────────────────────────────────────────────────────────
(Intercept)     11.473 (0.231) 49.57 153.8 <.001 *** [11.016, 11.930]
ses              2.385 (0.118) 20.24 157.8 <.001 *** [ 2.153,  2.618]
sectorCatholic   2.541 (0.345)  7.37 151.3 <.001 *** [ 1.860,  3.222]
─────────────────────────────────────────────────────────────────────
'df' is estimated by Satterthwaite approximation.

Standardized Coefficients (β):
Outcome Variable: mathach
──────────────────────────────────────────────────────────────────
                    β    S.E.     t    df     p      [95% CI of β]
──────────────────────────────────────────────────────────────────
ses             0.270 (0.013) 20.24 157.8 <.001 *** [0.244, 0.297]
sectorCatholic  0.185 (0.025)  7.37 151.3 <.001 *** [0.135, 0.234]
──────────────────────────────────────────────────────────────────

Random Effects:
──────────────────────────────────────────
 Cluster  K   Parameter   Variance     ICC
──────────────────────────────────────────
 id       160 (Intercept)  3.96456 0.09725
              ses          0.43431
 Residual                 36.80079
──────────────────────────────────────────

ANOVA-like table for random-effects: Single term deletions

Model:
mathach ~ ses + sector + (ses | id)
                  npar logLik   AIC    LRT Df Pr(>Chisq)
<none>               7 -23301 46616
```

```
ses in (ses | id)     5 -23306 46621 9.2963  2   0.009579 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Model 5-1. Context effect model

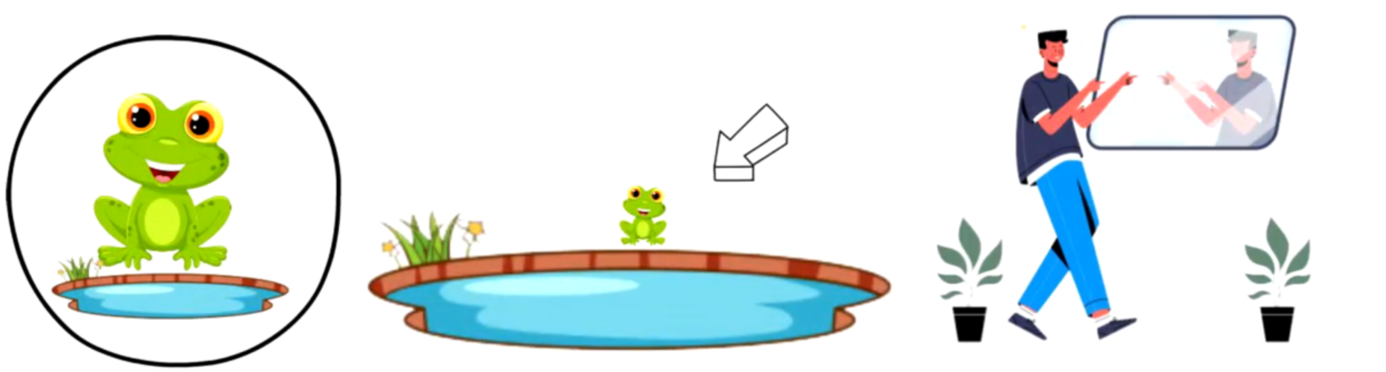
2. Context effect model

- Context effect: a group-level effect above and beyond the individual level effect
    - β_w: Difference in Y between two student *within* same school
    - β_b: Difference in mean of Y ($\bar{Y}$) *between* two schools
    - $\beta_c = \beta_b - \beta_w$: Difference in between two students who have the same individual SES, but who attend schools that differ by one unit of mean SES



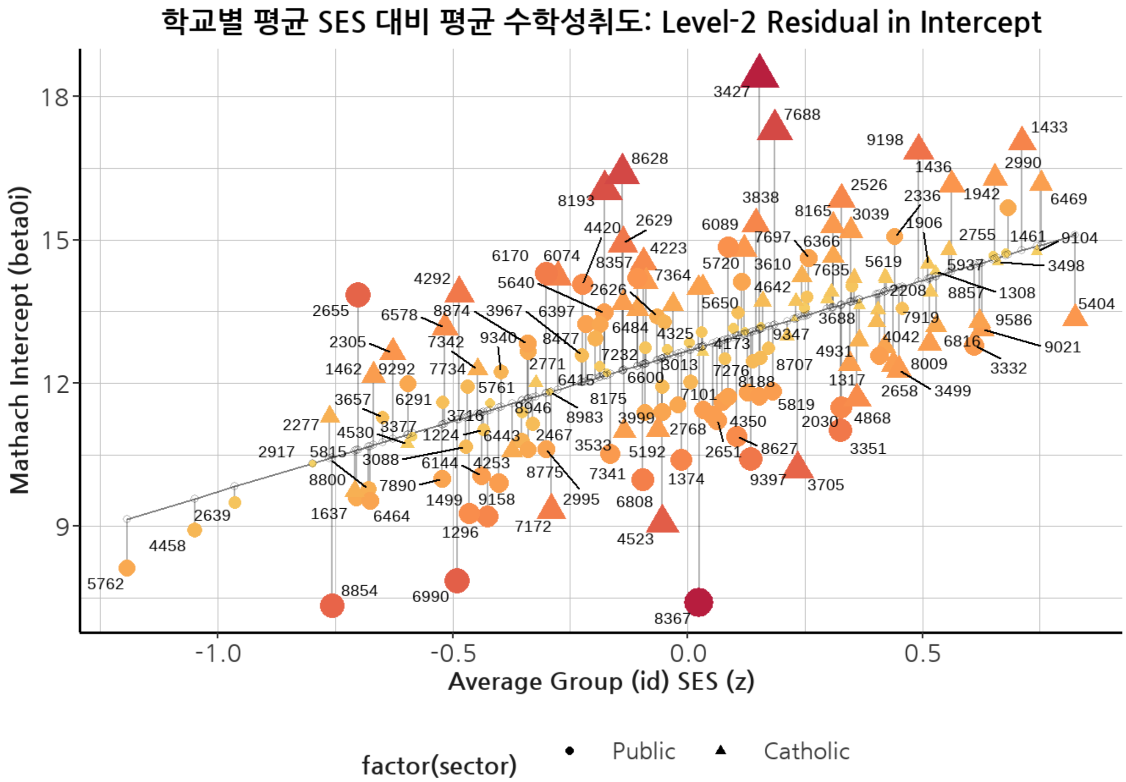
Model 5-1. Context effect model (Cont’d)

- 개구리-연못 이론 (Davis, 1966): Social Comparison theory에 기반
    - 조직연구에서 개인수준의 변수를 조직 또는 집단의 수준에서 aggregate 했을 때 가중되는 효과
    - 지금까지는 학생들의 SES의 수학성취도에 대한 영향을 Level-1 equation에서 학교의 영향과 상관없이 고려하였으나, 실제는 어떤 학교에서 학습은 영향 (e.g. Peer Effects, Proxy for other variable not in model)



Model 5-1. Context effect model (Cont’d)

```
Warning: ggrepel: 28 unlabeled data points (too many overlaps). Consider increasing max.overlaps
```



Model 5-1. Context effect model (Cont'd)

- 일반적으로 Group-mean Centering을 적용함
  - Level 1: $Y_{ij} = \beta_{0j} + \beta_{1j}(X_{ij} - \bar{X}_j) + \epsilon_{ij} \quad \epsilon_{ij} \sim N(0, \sigma^2)$
  - Level 2 intercept: $\beta_{0i} = \gamma_{00} + \gamma_{01}\bar{X}_j + u_{0j} \quad u_{0j} \sim N(0, \tau_{00})$
  - Level 2 slope: $\beta_{1j} = \gamma_{10} + u_{1j} \quad u_{1j} \sim N(0, \tau_{11})$
- Group-mean Centering을 선택하면 X_{ij}와 Y_{ij}의 관계는 집단 내와 집단 간 영향으로 구분되어지는데, 이를 분리하기 위한 contextual effect를 추가하여 완전한 집단 내 영향으로 추정

```
## Contextual Effect
model7 <- lmer(mathach ~ ses_groupmc + meanses + (1 | id), data = data_merged, REML = TRUE)
icc(model7)
```

```
# Intraclass Correlation Coefficient

    Adjusted ICC: 0.068
  Unadjusted ICC: 0.056
```

```
summary(model7)
```

```
Linear mixed model fit by REML. t-tests use Satterthwaite's method ['lmerModLmerTest']
Formula: mathach ~ ses_groupmc + meanses + (1 | id)
   Data: data_merged

REML criterion at convergence: 46568.6

Scaled residuals:
    Min      1Q  Median      3Q     Max
-3.1666 -0.7254  0.0174  0.7558  2.9454

Random effects:
 Groups   Name        Variance Std.Dev.
 id       (Intercept)  2.693   1.641
 Residual             37.019   6.084
Number of obs: 7185, groups:  id, 160

Fixed effects:
             Estimate Std. Error        df t value Pr(>|t|)
(Intercept)   12.6833     0.1494  153.6518   84.91   <2e-16 ***
ses_groupmc    2.1912     0.1087 7021.5092   20.16   <2e-16 ***
meanses        5.8662     0.3617  153.3666   16.22   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Correlation of Fixed Effects:
            (Intr) ss_grp
ses_groupmc 0.000
meanses     0.010  0.000
```

```
# the individual random effects (level 2 residuals of the intercept, i.e. the υ0i)
ranef(model7)
```

```
$id
     (Intercept)
1224 -0.32470863
1288  0.07365742
1296 -1.98322928
1308  0.28140400
1317 -1.19042740
1358 -0.93376699
1374 -1.93200718
1433  2.05275134
1436  1.61998267
1461  0.13077685
1462  1.40112873
1477  0.49848963
1499 -1.81811517
1637 -1.00791086
1906  0.23709066
1909  0.19099801
1942  0.96795426
1946  0.14887808
2030 -1.95854043
2208  0.19443542
2277  0.88456599
2305  1.77431117
2336  0.96248745
2458 -0.02729296
2467 -0.47574383
2526  1.97547043
2626  0.80310175
2629  2.44273319
2639 -0.30913112
2651 -1.44577342
2655  2.98789021
2658 -1.42405379
2755  0.41110188
2768 -0.95613325
2771  0.92169655
2818  0.41591664
2917 -0.01650581
2990  1.49558922
2995 -1.05274736
3013  0.14845208
3020  0.29354342
3039  1.35419720
3088 -0.56561826
3152  0.27179187
3332 -1.46287647
3351 -2.31499083
3377 -0.04820238
3427  4.79048715
3498 -0.12853090
3499 -1.50236903
3533 -1.15475088
3610  1.61913944
3657  0.50886785
3688 -0.30535475
3705 -2.85571159
3716  0.11469012
3838  2.01414101
3881 -1.11033363
3967  0.53174237
```

```
3992 -0.14417688
3999 -0.92786159
4042 -0.59768915
4173 -0.11151945
4223  1.90783563
4253 -0.87168559
4292  2.50339144
4325  0.67087649
4350 -1.21633499
4383 -1.06263385
4410  0.16791713
4420  1.75311274
4458 -0.55627204
4511  1.09471582
4523 -3.10731229
4530 -0.10355588
4642  1.01071353
4868 -1.77410398
4931 -0.83443374
5192 -1.27432061
5404 -1.69819828
5619  0.22112297
5640  1.21876774
5650  1.11721955
5667 -1.11253830
5720  1.11794767
5761  0.27643469
5762 -0.98757582
5783 -0.37176210
5815 -0.91806029
5819 -1.26253733
5838  0.06106942
5937  0.18586356
6074  2.18323058
6089  1.68154829
6144 -1.19082754
6170  1.97509942
6291  0.34783542
6366  0.93447507
6397  1.12502338
6415  0.21950393
6443 -0.77968877
6464 -1.10205476
6469  1.09406507
6484  0.94424947
6578  1.87953670
6600 -0.53420910
6808 -2.15850169
6816 -0.99821260
6897  0.28408658
6990 -3.04416888
7011  0.92760588
7101 -0.58646918
7172 -2.22203892
7232  0.30683273
7276 -0.36747427
7332  0.16366394
7341 -1.51751918
7342  0.89741072
7345 -1.23627997
7364  1.53378502
7635  0.75367687
```

```
7688  3.70518306
7697  1.06566914
7734  0.38751895
7890 -1.00499514
7919 -0.37776604
8009 -1.25589830
8150  0.28557284
8165  1.52713312
8175  0.01526995
8188 -0.51960683
8193  3.47412578
8202 -1.05806271
8357  1.54235512
8367 -4.17774372
8477  0.72130898
8531 -1.15970455
8627 -1.91719945
8628  3.80709702
8707 -0.55141639
8775 -0.95128440
8800 -0.83012371
8854 -2.80089462
8857 -0.34581005
8874  0.99194623
8946 -0.19931711
8983  0.02572273
9021 -1.33490468
9104 -0.16994877
9158 -1.41260481
9198  2.44045788
9225  0.36008266
9292  0.62814464
9340  0.56029937
9347 -0.31108743
9359  0.40514687
9397 -2.40973997
9508  1.22117550
9550 -1.29250630
9586 -1.18683318

with conditional variances for "id"
```

```
ranova(model7)
```

```
ANOVA-like table for random-effects: Single term deletions

Model:
mathach ~ ses_groupmc + meanses + (1 | id)
         npar logLik   AIC    LRT Df Pr(>Chisq)
<none>      5 -23284 46579
(1 | id)    4 -23417 46842 265.12  1  < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
## Use BruceR package
HLM_summary(model7,test.rand = T, digits = 3)
```

```
Hierarchical Linear Model (HLM)
(also known as) Linear Mixed Model (LMM)
(also known as) Multilevel Linear Model (MLM)

Model Information:
Formula: mathach ~ ses_groupmc + meanses + (1 | id)
Level-1 Observations: N = 7185
Level-2 Groups/Clusters: id, 160

Model Fit:
AIC = 46578.584
BIC = 46612.983
R_(m)² = 0.16733  (Marginal R²: fixed effects)
R_(c)² = 0.22379  (Conditional R²: fixed + random effects)
Omega² = 0.23072  (= 1 - proportion of unexplained variance)

ANOVA Table:
─────────────────────────────────────────────────────────────
               Sum Sq  Mean Sq NumDF   DenDF      F     p
─────────────────────────────────────────────────────────────
ses_groupmc  15051.53 15051.53  1.00 7021.51 406.59 <.001 ***
meanses       9737.33  9737.33  1.00  153.37 263.04 <.001 ***
─────────────────────────────────────────────────────────────

Fixed Effects:
Unstandardized Coefficients (b or γ):
Outcome Variable: mathach
───────────────────────────────────────────────────────────────────
                b/γ    S.E.     t     df     p      [95% CI of b/γ]
───────────────────────────────────────────────────────────────────
(Intercept)  12.683 (0.149) 84.91  153.7 <.001 *** [12.388, 12.978]
ses_groupmc   2.191 (0.109) 20.16 7021.5 <.001 *** [ 1.978,  2.404]
meanses       5.866 (0.362) 16.22  153.4 <.001 *** [ 5.152,  6.581]
───────────────────────────────────────────────────────────────────
'df' is estimated by Satterthwaite approximation.

Standardized Coefficients (β):
Outcome Variable: mathach
────────────────────────────────────────────────────────────────
                 β    S.E.     t     df     p      [95% CI of β]
────────────────────────────────────────────────────────────────
ses_groupmc  0.210 (0.010) 20.16 7021.5 <.001 *** [0.190, 0.231]
meanses      0.353 (0.022) 16.22  153.4 <.001 *** [0.310, 0.396]
────────────────────────────────────────────────────────────────

Random Effects:
──────────────────────────────────────────
 Cluster  K   Parameter   Variance     ICC
──────────────────────────────────────────
 id       160 (Intercept)  2.69253 0.06780
 Residual                 37.01906
──────────────────────────────────────────

ANOVA-like table for random-effects: Single term deletions

Model:
mathach ~ ses_groupmc + meanses + (1 | id)
         npar logLik   AIC    LRT Df Pr(>Chisq)
<none>      5 -23284 46579
(1 | id)    4 -23417 46842 265.12  1  < 2.2e-16 ***
```

```
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# LRtest
anova(model3, model7)
```

```
refitting model(s) with ML (instead of REML)
```

```
Data: data_merged
Models:
model3: mathach ~ ses_grandmc + (1 | id)
model7: mathach ~ ses_groupmc + meanses + (1 | id)
       npar   AIC   BIC logLik deviance  Chisq Df Pr(>Chisq)
model3    4 46649 46677 -23321    46641
model7    5 46574 46608 -23282    46564 77.195  1  < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Model 5-1. Context effect model (Cont'd)

```
Model Information:
Formula: mathach ~ ses_groupmc + ses_grouped + (1 ¦ id)
Level-1 Observations: N = 7185
Level-2 Groups/Clusters: id, 160

Model Fit:
AIC = 46578.584
BIC = 46612.983
R_(m)² = 0.16733  (Marginal R²: fixed effects)
R_(c)² = 0.22379  (Conditional R²: fixed + random effects)
Omega² = 0.23072  (= 1 - proportion of unexplained variance)

ANOVA Table:
────────────────────────────────────────────────────────────
               Sum Sq  Mean Sq NumDF   DenDF      F     p
────────────────────────────────────────────────────────────
ses_groupmc  15051.53 15051.53  1.00 7021.51 406.59 <.001 ***
ses_grouped   9737.33  9737.33  1.00  153.37 263.04 <.001 ***
────────────────────────────────────────────────────────────

Fixed Effects:
Unstandardized Coefficients (b or γ):
Outcome Variable: mathach
────────────────────────────────────────────────────────────────
                b/γ    S.E.     t     df     p      [95% CI of b/γ]
────────────────────────────────────────────────────────────────
(Intercept)  12.683 (0.149) 84.91  153.7 <.001 *** [12.388, 12.978]
ses_groupmc   2.191 (0.109) 20.16 7021.5 <.001 *** [ 1.978,  2.404]
ses_grouped   5.866 (0.362) 16.22  153.4 <.001 *** [ 5.152,  6.581]
────────────────────────────────────────────────────────────────
'df' is estimated by Satterthwaite approximation.
```

```
Standardized Coefficients (β):
Outcome Variable: mathach
────────────────────────────────────────────────────────────
                  β    S.E.      t     df     p       [95% CI of β]
────────────────────────────────────────────────────────────
ses_groupmc  0.210 (0.010) 20.16 7021.5 <.001 *** [0.190, 0.231]
ses_grouped  0.353 (0.022) 16.22  153.4 <.001 *** [0.310, 0.396]
────────────────────────────────────────────────────────────

Random Effects:
────────────────────────────────────────
 Cluster  K   Parameter   Variance     ICC
────────────────────────────────────────
 id      160 (Intercept)  2.69253 0.06780
 Residual                37.01906
────────────────────────────────────────

ANOVA-like table for random-effects: Single term deletions

Model:
mathach ~ ses_groupmc + ses_grouped + (1 ¦ id)
          npar  logLik    AIC    LRT Df Pr(>Chisq)
<none>       5 -23284 46579
(1 ¦ id)     4 -23417 46842 265.12  1  < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

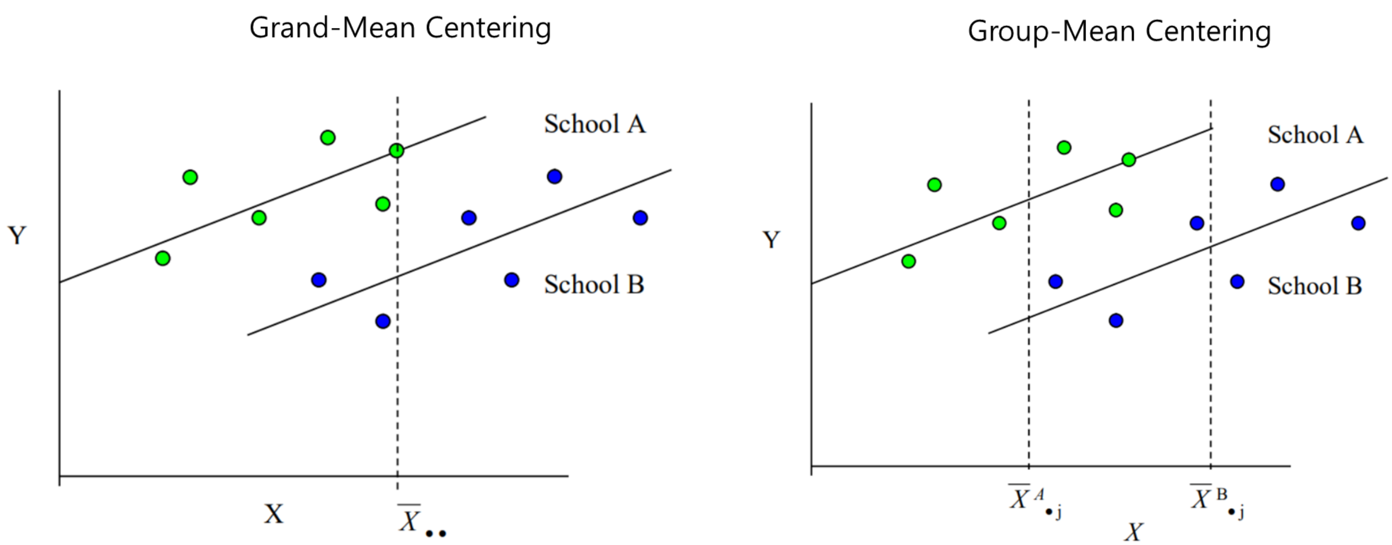
Module IV: Centering and Cross-level Interaction

Centering

- 일반적인 회귀분석에서 변수를 중심화하는 경우는 대부분 상호작용 효과를 분석할 때이다. 그리고 이 때 변수를 중심화하는 것은 절편(intercept)을 의미있는 값으로 해석하기 위해서, 그리고 다중공선성(multicollinearity)의 문제를 해결하기 위해서 사용됨
- Raudenbush and Bryk(2002): 변수 중심화는 1) 절편의 의미, 2) 독립변수 추정치의 의미, 3) 추정치의 수치적 안정성(numerical stability)에 영향

- Centering을 하지 않은 모형: $Y_{ij} = \beta_{0j} + \beta_{1j}X_{ij} + r_{ij}$
- e.g. X_{ij} = previous math achievement, 절편은 0일때의 기대값, 현실적으로 해당 변수가 0이 될 수 있는가?
- 변수 중심화의 두가지 유형
    - Grand mean centering (전체 평균 중심화): $Y_{ij} = \beta_{0j} + \beta_{1j}(X_{ij} - \bar{X}_{..}) + \epsilon_{ij}$
        - when expected value of Y_{ij} when X_{ij} = $\bar{X}_{..}$
        - β_{ij} (절편) = The mathach of the “average student” in school j
    - Group mean centering (집단 평균 중심화): $Y_{ij} = \beta_{0j} + \beta_{1j}(X_{ij} - \bar{X}_{.j}) + \epsilon_{ij}$
        - when expected value of Y_{ij} when X_{ij} = $\bar{X}_{.j}$
        - β_{ij} (절편) = The mathach of the “average student” in the all sample of schools
        - 어느 학교에서든 $\bar{X}_{..}$ 인 값들은 서로 같음

Grand-mean Centering & Group-mean Centering



Grand-mean Centering & Group-mean Centering (Cont’d)

- 그래서 도대체 언제 Centering을 해야 하는가? 출처 (https://ukchanoh.wordpress.com/2015/07/08/multilevel-centering/)

    1. **Level-1 변수들 간의 관계에 관심이 있을 때**
    - 개인-집단의 2수준 자료에서 X_ij에는 집단 내 변량(within-group variation)과 집단 간 변량(between-group variation)이 혼재됨 (Contextual Effect Model 참고)
    - 중심화를 하지 않거나, 전체평균 중심화를 하면 X_{1j}의 계수인 γ_{10}은 집단 내 효과와 집단 간 효과의 가중평균이 되므로, 집단 내 효과도 아니고 집단 간 효과도 아니고 총 효과도 아닌 값으로 추정이 이루어짐
    - (만약 집단 평균 중심화를 하지 않는다면 β_w = β_b를 가정하므로 이게 현실적인 가정인지 고민이 필요함함)
    - 따라서, **집단평균 중심화**를 통해 집단 간(between) 변량을 제거하여 순수한 β_{Within} 값을 추정해야 함
    - 2 수준에 집단평균 변수 $\bar{X}_{.j}$를 투입하여 Contextual Effect를 함께 고려하는 것이 바람직하지만, 그렇지 않은 논문들도 多
    2. **Level-2 변수들과과의 관계에 관심이 있을 때**
    - Level-1 변수를 집단평균 중심화하면 Level-1 절편은 조정되지 않은 집단의 평균값이기 때문에 집단 간 변량이 제거되어 Level-2에서 영향을 미치지 않음
    - Level-1 변수를 전체평균 중심화하면 Level-1 절편은 조정된 평균값이기 때문에 Level-2 변수의 영향을 추정할 때 Level-1 변수의 집단 간 평균값 차이에 의해 조정
    - 따라서, Level-1 변수를 **중심화하지 않거**나, **전체평균 중심화**를 해야함

Grand-mean Centering & Group-mean Centering (Cont’d)

- 그래서 도대체 언제 Centering을 해야 하는가? 출처 (https://ukchanoh.wordpress.com/2015/07/08/multilevel-centering/)

    3. **Level-1 & Level-2 수준 간 상호작용에 관심이 있을 때**
    - Level-1 변수(X_{ij})의 영향이 Level-2 변수(Z_j)에 따라 어떻게 달라지는지가 연구문제일 때

- Level-1 변수가 집단 내 효과(β_w를 의미하도록 **집단평균 중심화**를 하는 것이 적절 (Hofmann and Gavin, 1998)
- 전체평균 중심화 하면 집단 간 상호작용과 수준 간 상호작용의 혼재가 발생

Grand-mean Centering & Group-mean Centering (Cont'd)

```
# id는 각 학교들의 식별자이기에, 식별자를 기준으로 집단화

data_merged <- data_merged %>%
  # 전체평균 중심화 (ses - grandmean)
  mutate(ses_grandmc = ses - mean(ses))%>%
  # 맥락효과(meanses = groupmean)와 집단평균 중심화 (ses_groupmc = ses - groupmean)
  # id 기준으로 observations들 group화
  group_by(id) %>%
  mutate(meanses = mean(ses),
         ses_groupmc = ses - meanses) %>%
  # group화 해제 (반드시 해야함)
  ungroup()

# 각 학교별 인원, 평균 ses 확인하는 법
data_merged %>%
  group_by(id) %>%
  # n = 인원수, meanses = 평균 ses 점수
  summarise(n = n(),
            meanses = mean(ses))
```

Model 5-2. Slopes-as-Outcomes Model

- Slopes-as-Outcomes Model의 경우 Cross-level Interactions model로 해석되기도 함
- conditional models because intercepts and slopes are conditioned by level-2 predictors (따라서, **집단평균 중심화**)
  - Estimation techniques similar to previous models except estimates are conditioned
- 목적: What level-2 factors predict differences in slopes between level-2 groups?
  - RQ: (e.g. whether variation in SES effects across schools can be attributed to the type of school – public vs Catholic)
- 모형
  - Level 1 Model: $Y_{ij} = \beta_{0j} + \beta_{1j}X_{ij} + \epsilon \quad \epsilon \sim N(0, \sigma^2)$
  - Level 2 Model (intercept): $\beta_{0j} = \gamma_{00} + \gamma_{01}Z_j + u_{0j} \quad u_{0j} \sim N(0, \tau_{00}^2)$
  - Level 2 Model (slope): $\beta_{1j} = \gamma_{10} + \gamma 11 Z_j + u_{1j} \quad u_{1j} \sim N(0, \tau_{11}^2)$
- Overall Model Example: (intercept + fixed + random)

$$MATHCH_{ij} = \underline{\beta_{0j} + \gamma_{00}} + \underline{\gamma_{01}(SECTOR_j) + \gamma_{10}(SES_{ij}) + \gamma_{11}(SECTOR_j)(SES_{ij})} + \underline{u_{1j}(SES_{ij}) + u_{0j} + \epsilon_{ij}}$$

- 장점: allows us to explain the variation in both intercepts and slopes
- HLM의 최종 목적인 모형이며, 가장 어려운 모형이기도 하다.

Model 5-2. Slopes-as-Outcomes Model (Cont'd)

- Random coefficient included

```
## with random coefficient
model8 <- lmer(mathach ~ ses_groupmc * sector + (ses_groupmc | id), data = data_merged, REML = FALSE)

summary(model8)
```

```
Linear mixed model fit by maximum likelihood . t-tests use Satterthwaite's method ['lmerModLmerTest']
Formula: mathach ~ ses_groupmc * sector + (ses_groupmc | id)
   Data: data_merged

     AIC      BIC   logLik deviance df.resid
 46649.9  46704.9 -23316.9  46633.9     7177

Scaled residuals:
     Min       1Q   Median       3Q      Max
-3.06465 -0.73019  0.01579  0.75355  2.94165

Random effects:
 Groups   Name        Variance Std.Dev. Corr
 id       (Intercept)  6.6404  2.5769
          ses_groupmc  0.2399  0.4898   0.82
 Residual             36.7055  6.0585
Number of obs: 7185, groups:  id, 160

Fixed effects:
                           Estimate Std. Error       df t value Pr(>|t|)
(Intercept)                 11.3939     0.2909 160.4371  39.168  < 2e-16 ***
ses_groupmc                  2.8029     0.1539 142.9981  18.213  < 2e-16 ***
sectorCatholic               2.8075     0.4363 155.5873   6.434 1.46e-09 ***
ses_groupmc:sectorCatholic  -1.3414     0.2322 153.4473  -5.776 4.12e-08 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Correlation of Fixed Effects:
            (Intr) ss_grp sctrCt
ses_groupmc  0.260
sectorCthlc -0.667 -0.173
ss_grpmc:sC -0.172 -0.663  0.262
```

```
icc(model8)
```

```
# Intraclass Correlation Coefficient

    Adjusted ICC: 0.155
  Unadjusted ICC: 0.141
```

```
# the individual random effects (level 2 residuals of the intercept, i.e. the u0i)
ranef(model8)
```

```
$id
      (Intercept)  ses_groupmc
1224 -1.504882175 -0.237296204
1288  1.743302075  0.276370154
1296 -3.435475425 -0.584390948
1308  1.567430353  0.230355599
1317 -0.919772745 -0.145009596
1358 -0.036542025  0.053074651
1374 -1.315680670 -0.171344553
1433  4.749429586  0.733321180
1436  3.456583255  0.528417036
1461  4.830283474  0.840003051
1462 -3.457173270 -0.607130772
1477  0.010744350 -0.011282911
1499 -3.288900094 -0.440731266
1637 -3.566200995 -0.530218676
1906  1.635027440  0.272517924
1909  2.508888588  0.382361916
1942  5.518415360  0.804376893
1946  1.362199861  0.234366278
2030  0.523189726  0.018211947
2208  1.149748488  0.222267049
2277 -4.658199823 -0.873913345
2305 -2.933042345 -0.559544961
2336  4.498381220  0.641291406
2458 -0.109376596  0.056941915
2467 -1.108711847 -0.159242298
2526  2.514864773  0.321032063
2626  1.789754713  0.304044801
2629  0.556567084  0.012786707
2639 -4.360043947 -0.769781677
2651 -0.158442700  0.041300518
2655  0.978455323  0.245439017
2658 -0.648910172 -0.054603682
2755  1.972622243  0.263191153
2768 -0.342360291 -0.023321010
2771  0.458862937  0.112558739
2818  2.173530082  0.328530743
2917 -3.147994582 -0.572190520
2990  3.765527629  0.557408529
2995 -1.759045112 -0.351384056
3013  1.128844862  0.197671245
3020  0.186578421  0.037276109
3039  2.252461376  0.375514972
3088 -2.025148209 -0.350338110
3152  1.614281409  0.231535892
3332  2.467626671  0.355756806
3351  0.031019777 -0.014145697
3377 -2.166150727 -0.475868606
3427  4.798748656  0.633893163
3498  1.898213961  0.229728586
3499 -0.824774866 -0.139123857
3533 -3.448250193 -0.573259782
3610  1.132517012  0.243589056
3657 -1.596481535 -0.177990494
3688  0.403550454  0.063357088
3705 -3.432019538 -0.526360368
3716 -0.638744004  0.067348743
3838  1.621342338  0.198960477
3881  0.471163422  0.062117826
3967  0.601298658  0.110893991
```

3992	0.354054132	0.017286382
3999	-0.315725308	0.010717699
4042	0.116525957	0.029088670
4173	-1.202161769	-0.114052002
4223	0.421497222	0.097615864
4253	-4.446661871	-0.759201929
4292	-1.312763378	-0.282187828
4325	1.645725478	0.236428143
4350	0.480212176	0.119411053
4383	0.196850021	0.087951977
4410	1.816005825	0.273555134
4420	2.104776050	0.323648400
4458	-5.054473386	-0.824867204
4511	-0.782409447	-0.173124657
4523	-5.118807040	-0.717549404
4530	-4.683003933	-0.687351739
4642	2.939450228	0.459775696
4868	-1.621748621	-0.250697488
4931	-0.404901338	-0.089185499
5192	-3.118059184	-0.464993850
5404	1.091923027	0.158361067
5619	1.294477285	0.370646709
5640	1.645185012	0.286813179
5650	0.003402735	-0.041212934
5667	-0.275762128	0.058237115
5720	0.130769997	0.066084209
5761	-2.631913321	-0.304176901
5762	-6.248037404	-1.032589109
5783	1.481003738	0.234064308
5815	-3.346957374	-0.510833028
5819	0.626251669	0.064225953
5838	1.888095380	0.265116093
5937	4.420988552	0.644076286
6074	-0.377690945	-0.053457262
6089	3.473439098	0.488708366
6144	-2.500433609	-0.374911333
6170	2.333438649	0.410289274
6291	-1.046635032	-0.128060657
6366	1.319706528	0.198654553
6397	1.265637315	0.181458166
6415	0.463605061	0.105218018
6443	-1.857759621	-0.404936606
6464	-3.673273738	-0.601880707
6469	3.862020277	0.589576910
6484	1.165505179	0.101834525
6578	-1.944723061	-0.250102187
6600	0.401636948	0.162517841
6808	-1.891572467	-0.308319628
6816	0.295914981	0.037855656
6897	3.346673512	0.536450885
6990	-5.006463044	-0.862371768
7011	-0.154951403	0.068448906
7101	0.255860690	-0.016906045
7172	-5.427203330	-0.832637131
7232	1.127596021	0.245825541
7276	1.227233712	0.241875711
7332	0.444511122	0.109157135
7341	-1.523575491	-0.299630731
7342	-2.771568613	-0.433597863
7345	0.072919882	0.114336236
7364	-0.066317181	-0.037485751
7635	0.831165626	0.169363855

```
7688  3.746413256  0.518571863
7697  3.678984960  0.568937785
7734 -0.366336583  0.054995433
7890 -2.895220531 -0.560457205
7919  3.033439869  0.489241773
8009 -0.099698540 -0.012249433
8150  0.493707675  0.020238638
8165  2.023087317  0.317026187
8175  0.186727491 -0.009436250
8188  1.217948715  0.229071441
8193  1.825160752  0.302002641
8202  0.328521476  0.080709097
8357  2.417372889  0.349924718
8367 -5.008049560 -0.809815554
8477  1.050620226  0.203917123
8531  1.894400069  0.304307994
8627 -0.518145969 -0.125561466
8628  2.108241677  0.306409158
8707  1.367439403  0.236341222
8775 -1.817236969 -0.348806014
8800 -5.709618505 -0.818375323
8854 -6.093321604 -0.948376904
8857  0.977104978  0.123050584
8874  0.650169810  0.144989117
8946 -1.001067708 -0.217540939
8983 -0.435346970 -0.123836968
9021  0.496197212  0.115428360
9104  2.381150459  0.364326044
9158 -2.474312105 -0.302870910
9198  4.177101642  0.665623146
9225  2.813442948  0.425851265
9292 -0.921393879 -0.163340253
9340 -0.159244162 -0.014750036
9347 -0.522595783 -0.012536159
9359  0.861451106  0.049729275
9397 -0.939294159 -0.153984032
9508 -0.434540628 -0.009205333
9550 -0.171975201  0.012487393
9586  0.611077182  0.100268415

with conditional variances for "id"
```

```
ranova(model8)
```

```
ANOVA-like table for random-effects: Single term deletions

Model:
mathach ~ ses_groupmc + sector + (ses_groupmc | id) + ses_groupmc:sector
                                  npar logLik   AIC   LRT Df Pr(>Chisq)   
<none>                               8 -23317 46650                       
ses_groupmc in (ses_groupmc | id)    6 -23323 46658 11.83  2   0.002699 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
## Use BruceR package
HLM_summary(model8, test.rand = T, digits = 3)
```

```
Hierarchical Linear Model (HLM)
(also known as) Linear Mixed Model (LMM)
(also known as) Multilevel Linear Model (MLM)

Model Information:
Formula: mathach ~ ses_groupmc * sector + (ses_groupmc | id)
Level-1 Observations: N = 7185
Level-2 Groups/Clusters: id, 160

Model Fit:
AIC = 46649.881
BIC = 46704.919
R_(m)² = 0.09015  (Marginal R²: fixed effects)
R_(c)² = 0.23139  (Conditional R²: fixed + random effects)
Omega² = 0.24000  (= 1 - proportion of unexplained variance)

ANOVA Table:
───────────────────────────────────────────────────────────────────
                      Sum Sq  Mean Sq NumDF  DenDF      F     p
───────────────────────────────────────────────────────────────────
ses_groupmc         12378.08 12378.08  1.00 153.45 337.23 <.001 ***
sector               1519.54  1519.54  1.00 155.59  41.40 <.001 ***
ses_groupmc:sector   1224.67  1224.67  1.00 153.45  33.36 <.001 ***
───────────────────────────────────────────────────────────────────

Fixed Effects:
Unstandardized Coefficients (b or γ):
Outcome Variable: mathach
─────────────────────────────────────────────────────────────────────────────────
                               b/γ    S.E.      t    df     p       [95% CI of b/γ]
─────────────────────────────────────────────────────────────────────────────────
(Intercept)                 11.394 (0.291) 39.17 160.4 <.001 *** [10.819, 11.968]
ses_groupmc                  2.803 (0.154) 18.21 143.0 <.001 *** [ 2.499,  3.107]
sectorCatholic               2.807 (0.436)  6.43 155.6 <.001 *** [ 1.946,  3.669]
ses_groupmc:sectorCatholic  -1.341 (0.232) -5.78 153.4 <.001 *** [-1.800, -0.883]
─────────────────────────────────────────────────────────────────────────────────
'df' is estimated by Satterthwaite approximation.

Standardized Coefficients (β):
Outcome Variable: mathach
─────────────────────────────────────────────────────────────────────────────────
                                 β    S.E.      t    df     p         [95% CI of β]
─────────────────────────────────────────────────────────────────────────────────
ses_groupmc                  0.269 (0.015) 18.21 143.0 <.001 *** [ 0.240,  0.298]
sectorCatholic               0.204 (0.032)  6.43 155.6 <.001 *** [ 0.141,  0.267]
ses_groupmc:sectorCatholic  -0.085 (0.015) -5.78 153.4 <.001 *** [-0.114, -0.056]
─────────────────────────────────────────────────────────────────────────────────

Random Effects:
──────────────────────────────────────────
 Cluster  K   Parameter   Variance     ICC
──────────────────────────────────────────
 id       160 (Intercept)  6.64044 0.15320
              ses_groupmc  0.23991
 Residual                 36.70554
──────────────────────────────────────────

ANOVA-like table for random-effects: Single term deletions

Model:
```

```
mathach ~ ses_groupmc + sector + (ses_groupmc | id) + ses_groupmc:sector
                                  npar logLik   AIC    LRT Df Pr(>Chisq)
<none>                               8 -23317 46650
ses_groupmc in (ses_groupmc | id)    6 -23323 46658 11.83  2   0.002699 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

- Random coefficient not included

```
## without random coefficient
model9 <- lmer(mathach ~ ses_groupmc * sector + (1 | id), data = data_merged, REML = FALSE)

summary(model9)
```

```
Linear mixed model fit by maximum likelihood . t-tests use Satterthwaite's method ['lmerModLmerTest']
Formula: mathach ~ ses_groupmc * sector + (1 | id)
   Data: data_merged

     AIC      BIC   logLik deviance df.resid
 46657.7  46699.0 -23322.9  46645.7     7179

Scaled residuals:
     Min       1Q   Median       3Q      Max
-3.04691 -0.72775  0.01723  0.75634  3.00238

Random effects:
 Groups   Name        Variance Std.Dev.
 id       (Intercept)  6.645   2.578
 Residual             36.808   6.067
Number of obs: 7185, groups:  id, 160

Fixed effects:
                            Estimate Std. Error        df t value Pr(>|t|)
(Intercept)                  11.3930     0.2910  160.4818  39.145  < 2e-16 ***
ses_groupmc                   2.7821     0.1446 7021.9891  19.245  < 2e-16 ***
sectorCatholic                2.8052     0.4366  155.6151   6.426 1.52e-09 ***
ses_groupmc:sectorCatholic   -1.3486     0.2184 7021.9891  -6.175 6.98e-10 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Correlation of Fixed Effects:
            (Intr) ss_grp sctrCt
ses_groupmc  0.000
sectorCthlc -0.667  0.000
ss_grpmc:sC  0.000 -0.662  0.000
```

```
icc(model9)
```

```
# Intraclass Correlation Coefficient

    Adjusted ICC: 0.153
  Unadjusted ICC: 0.139
```

```
# the individual random effects (level 2 residuals of the intercept, i.e. the u0i)
ranef(model9)
```

```
$id
     (Intercept)
1224 -1.50067378
1288  1.73366743
1296 -3.36831221
1308  1.61108690
1317 -0.91491393
1358 -0.15764794
1374 -1.38960989
1433  4.76655447
1436  3.47583190
1461  4.66634454
1462 -3.37466504
1477  0.02779710
1499 -3.37942254
1637 -3.62513139
1906  1.61607428
1909  2.52983532
1942  5.64048714
1946  1.32696692
2030  0.61295740
2208  1.10450706
2277 -4.49260483
2305 -2.82671789
2336  4.58439291
2458 -0.19368040
2467 -1.12556732
2526  2.60194742
2626  1.74869851
2629  0.64673326
2639 -4.22082391
2651 -0.26940275
2655  0.86051385
2658 -0.71412383
2755  2.03811090
2768 -0.41427630
2771  0.40984049
2818  2.19088384
2917 -3.02441883
2990  3.81003270
2995 -1.64837975
3013  1.10259843
3020  0.18016831
3039  2.18840600
3088 -1.96766359
3152  1.64121313
3332  2.51809287
3351  0.06321059
3377 -1.96445743
3427  4.95701514
3498  1.98481713
3499 -0.80439962
3533 -3.39710273
3610  1.06461924
3657 -1.68842356
3688  0.40579279
3705 -3.44270567
3716 -0.90240915
3838  1.69114618
3881  0.49002271
3967  0.58027130
```

```
3992  0.40472003
3999 -0.40069450
4042  0.10789598
4173 -1.30875987
4223  0.37791437
4253 -4.36813495
4292 -1.22908773
4325  1.67223147
4350  0.39596600
4383  0.05950448
4410  1.83241100
4420  2.11503681
4458 -5.00409695
4511 -0.72035338
4523 -5.23002890
4530 -4.72686651
4642  2.93913757
4868 -1.62350200
4931 -0.37186117
5192 -3.16294000
5404  1.10901753
5619  1.12373890
5640  1.61059215
5650  0.06708782
5667 -0.38499591
5720  0.07615523
5761 -2.76552443
5762 -6.14772198
5783  1.47550945
5815 -3.37402272
5819  0.67151319
5838  1.94845045
5937  4.51965255
6074 -0.38137293
6089  3.57626644
6144 -2.52290914
6170  2.20612020
6291 -1.10999738
6366  1.33108111
6397  1.28451753
6415  0.42374440
6443 -1.61858821
6464 -3.61151561
6469  3.88042064
6484  1.31179232
6578 -2.00577834
6600  0.28291650
6808 -1.87128857
6816  0.30893431
6897  3.32837036
6990 -4.90367694
7011 -0.32932948
7101  0.38146914
7172 -5.44576661
7232  1.03896536
7276  1.16467623
7332  0.39260813
7341 -1.44217068
7342 -2.76745985
7345 -0.04953696
7364 -0.02313776
7635  0.78236463
```

```
7688  3.83112261
7697  3.69002032
7734 -0.66572864
7890 -2.75288404
7919  3.00681776
8009 -0.10150076
8150  0.58102731
8165  2.02420178
8175  0.26124687
8188  1.13787048
8193  1.80193579
8202  0.27578888
8357  2.48004424
8367 -4.90100362
8477  0.98220252
8531  1.88148765
8627 -0.46108239
8628  2.13620052
8707  1.33668652
8775 -1.72671867
8800 -5.84963396
8854 -6.09765881
8857  1.01122508
8874  0.57375198
8946 -0.92916239
8983 -0.36170368
9021  0.45360378
9104  2.39291458
9158 -2.57833869
9198  4.15214896
9225  2.83769676
9292 -0.86203361
9340 -0.18004719
9347 -0.60102335
9359  0.97095366
9397 -0.92813247
9508 -0.53832879
9550 -0.25512078
9586  0.60838708

with conditional variances for "id"
```

```
ranova(model9)
```

```
ANOVA-like table for random-effects: Single term deletions

Model:
mathach ~ ses_groupmc + sector + (1 | id) + ses_groupmc:sector
         npar logLik   AIC   LRT Df Pr(>Chisq)
<none>      6 -23323 46658
(1 | id)    5 -23711 47432 776.2  1  < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
## Use BruceR package
HLM_summary(model9, test.rand = T, digits = 3)
```

```
Hierarchical Linear Model (HLM)
(also known as) Linear Mixed Model (LMM)
(also known as) Multilevel Linear Model (MLM)

Model Information:
Formula: mathach ~ ses_groupmc * sector + (1 | id)
Level-1 Observations: N = 7185
Level-2 Groups/Clusters: id, 160

Model Fit:
AIC = 46657.711
BIC = 46698.989
R_(m)² = 0.08924  (Marginal R²: fixed effects)
R_(c)² = 0.22851  (Conditional R²: fixed + random effects)
Omega² = 0.23717  (= 1 - proportion of unexplained variance)

ANOVA Table:
────────────────────────────────────────────────────────────────────
                      Sum Sq  Mean Sq NumDF   DenDF      F     p
────────────────────────────────────────────────────────────────────
ses_groupmc         13715.34 13715.34  1.00 7021.99 372.62 <.001 ***
sector               1519.86  1519.86  1.00  155.62  41.29 <.001 ***
ses_groupmc:sector   1403.55  1403.55  1.00 7021.99  38.13 <.001 ***
────────────────────────────────────────────────────────────────────

Fixed Effects:
Unstandardized Coefficients (b or γ):
Outcome Variable: mathach
──────────────────────────────────────────────────────────────────────────────────
                               b/γ    S.E.      t     df     p       [95% CI of b/γ]
──────────────────────────────────────────────────────────────────────────────────
(Intercept)                 11.393 (0.291) 39.15  160.5 <.001 *** [10.818, 11.968]
ses_groupmc                  2.782 (0.145) 19.24 7022.0 <.001 *** [ 2.499,  3.065]
sectorCatholic               2.805 (0.437)  6.43  155.6 <.001 *** [ 1.943,  3.668]
ses_groupmc:sectorCatholic  -1.349 (0.218) -6.18 7022.0 <.001 *** [-1.777, -0.920]
──────────────────────────────────────────────────────────────────────────────────
'df' is estimated by Satterthwaite approximation.

Standardized Coefficients (β):
Outcome Variable: mathach
──────────────────────────────────────────────────────────────────────────────────
                                 β    S.E.      t     df     p         [95% CI of β]
──────────────────────────────────────────────────────────────────────────────────
ses_groupmc                  0.267 (0.014) 19.24 7022.0 <.001 *** [ 0.240,  0.294]
sectorCatholic               0.204 (0.032)  6.43  155.6 <.001 *** [ 0.141,  0.267]
ses_groupmc:sectorCatholic  -0.086 (0.014) -6.18 7022.0 <.001 *** [-0.113, -0.059]
──────────────────────────────────────────────────────────────────────────────────

Random Effects:
──────────────────────────────────────────
 Cluster  K   Parameter   Variance     ICC
──────────────────────────────────────────
 id       160 (Intercept)  6.64476 0.15292
 Residual                 36.80827
──────────────────────────────────────────

ANOVA-like table for random-effects: Single term deletions

Model:
mathach ~ ses_groupmc + sector + (1 | id) + ses_groupmc:sector
```

```
         npar logLik   AIC   LRT Df Pr(>Chisq)    
<none>      6 -23323 46658                        
(1 | id)    5 -23711 47432 776.2  1  < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# LRtest
anova(model8, model9) # with randomness is better
```

```
Data: data_merged
Models:
model9: mathach ~ ses_groupmc * sector + (1 | id)
model8: mathach ~ ses_groupmc * sector + (ses_groupmc | id)
       npar   AIC   BIC logLik deviance Chisq Df Pr(>Chisq)   
model9    6 46658 46699 -23323    46646                       
model8    8 46650 46705 -23317    46634 11.83  2   0.002699 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Model 5-2. Slopes-as-Outcomes Model (Cont'd)

```
Model Information:
Formula: mathach ~ ses_groupmc * sector + (ses_groupmc ¦ id)
Level-1 Observations: N = 7185
Level-2 Groups/Clusters: id, 160

Model Fit:
AIC = 46649.881
BIC = 46704.919
R_(m)² = 0.09015  (Marginal R²: fixed effects)
R_(c)² = 0.23139  (Conditional R²: fixed + random effects)
Omega² = 0.24000  (= 1 - proportion of unexplained variance)

ANOVA Table:
─────────────────────────────────────────────────────────────────
                     Sum Sq  Mean Sq NumDF  DenDF      F     p
─────────────────────────────────────────────────────────────────
ses_groupmc        12378.08 12378.08  1.00 153.45 337.23 <.001 ***
sector              1519.54  1519.54  1.00 155.59  41.40 <.001 ***
ses_groupmc:sector  1224.67  1224.67  1.00 153.45  33.36 <.001 ***
─────────────────────────────────────────────────────────────────

Fixed Effects:
Unstandardized Coefficients (b or γ):
Outcome Variable: mathach
──────────────────────────────────────────────────────────────────────────────
                               b/γ    S.E.     t    df     p     [95% CI of b/γ]
──────────────────────────────────────────────────────────────────────────────
(Intercept)                 11.394 (0.291) 39.17 160.4 <.001 *** [10.819, 11.968]
ses_groupmc                  2.803 (0.154) 18.21 143.0 <.001 *** [ 2.499,  3.107]
sectorCatholic               2.807 (0.436)  6.43 155.6 <.001 *** [ 1.946,  3.669]
ses_groupmc:sectorCatholic  -1.341 (0.232) -5.78 153.4 <.001 *** [-1.800, -0.883]
──────────────────────────────────────────────────────────────────────────────
'df' is estimated by Satterthwaite approximation.
```

```
Standardized Coefficients (β):
Outcome Variable: mathach
──────────────────────────────────────────────────────────────────────────────
                                β    S.E.     t    df     p       [95% CI of β]
──────────────────────────────────────────────────────────────────────────────
ses_groupmc                 0.269 (0.015) 18.21 143.0 <.001 *** [ 0.240,  0.298]
sectorCatholic              0.204 (0.032)  6.43 155.6 <.001 *** [ 0.141,  0.267]
ses_groupmc:sectorCatholic -0.085 (0.015) -5.78 153.4 <.001 *** [-0.114, -0.056]
──────────────────────────────────────────────────────────────────────────────

Random Effects:
──────────────────────────────────────────
 Cluster  K   Parameter   Variance     ICC
──────────────────────────────────────────
 id       160 (Intercept)  6.64044 0.15320
              ses_groupmc  0.23991        
 Residual                 36.70554        
──────────────────────────────────────────

ANOVA-like table for random-effects: Single term deletions

Model:
mathach ~ ses_groupmc + sector + (ses_groupmc ¦ id) + ses_groupmc:sector
                               npar logLik   AIC   LRT Df Pr(>Chisq)   
<none>                            8 -23317 46650                       
ses_groupmc in (ses_groupmc ¦ id)    6 -23323 46658 11.83  2   0.002699 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

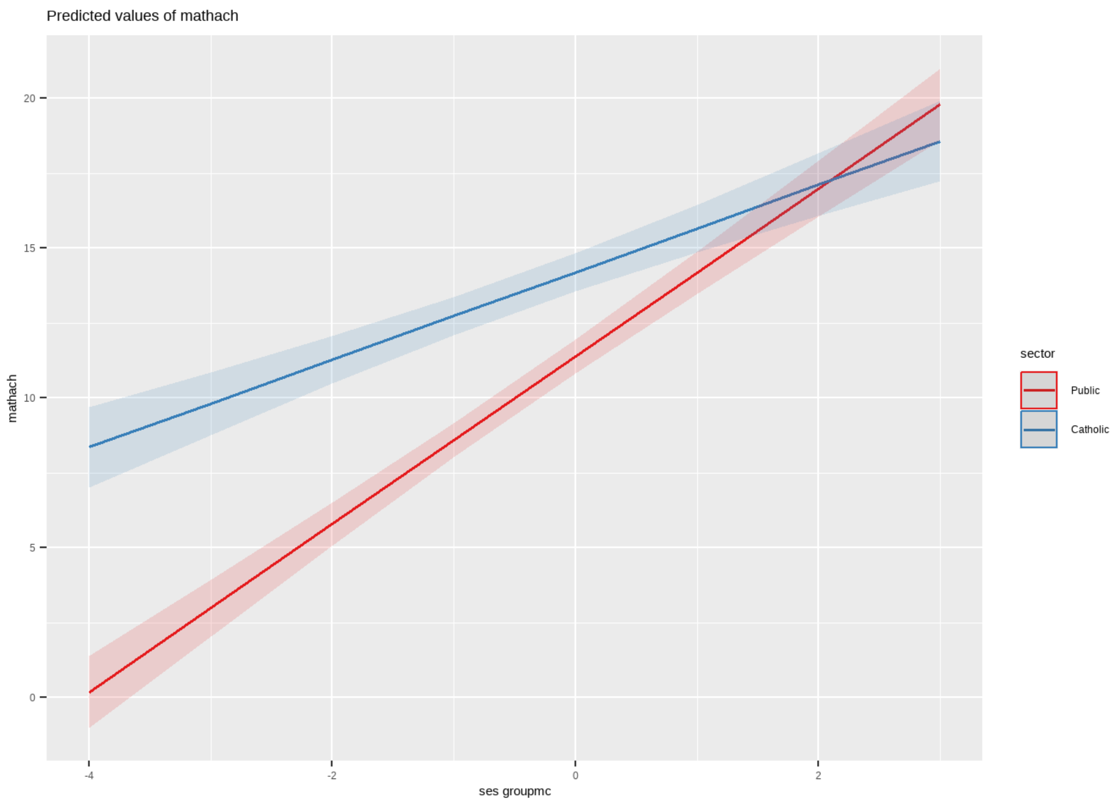
Model 5-2. Slopes-as-Outcomes: Marginsplot

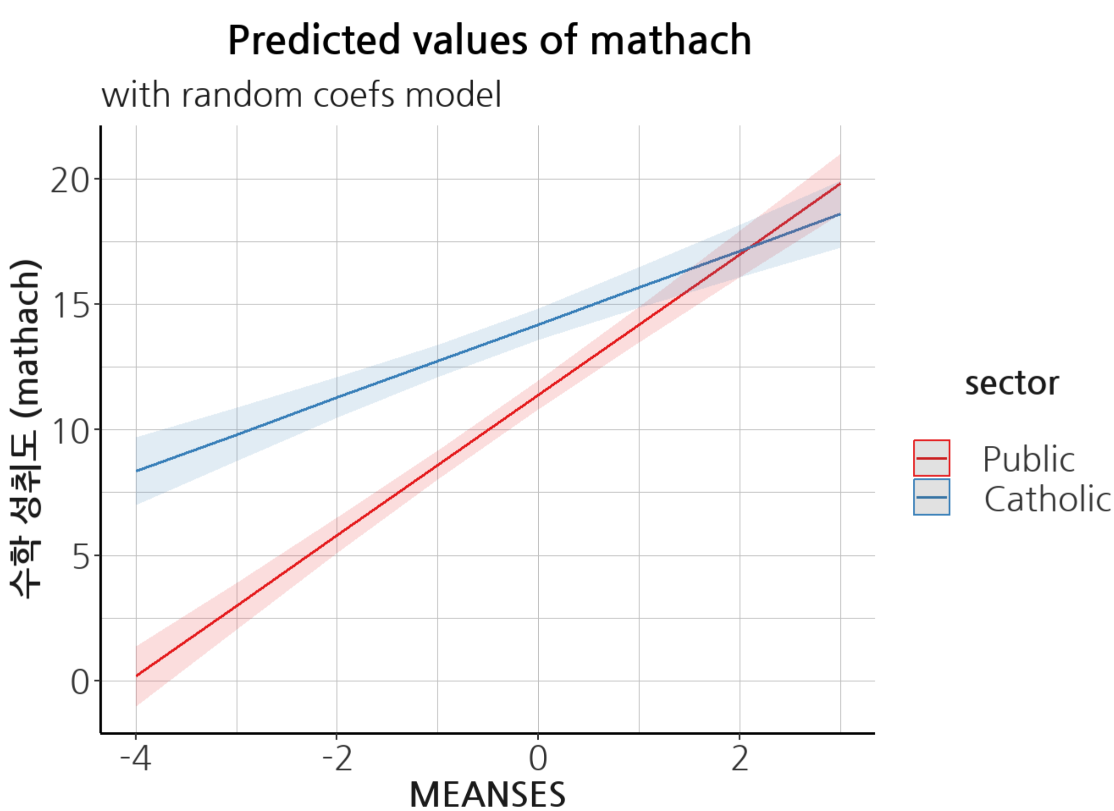
```
# Marginsplot
# install.packages("sjPlot")
# install.packages("effects")
library(sjPlot)
```

```
다음의 패키지를 부착합니다: 'sjPlot'
```

```
The following objects are masked from 'package:cowplot':

    plot_grid, save_plot
```

```
plot_model(model8, type = 'pred', terms = c('ses_groupmc','sector'), ci.lvl = 0.95)
```





Model Specification Summary

```
Please cite as:
```

```
Hlavac, Marek (2022). stargazer: Well-Formatted Regression and Summary Statistics Tables.
```

```
R package version 5.2.3. https://CRAN.R-project.org/package=stargazer
```

```
<table style="text-align:center"><caption><strong>Model comparison over Mathach ~ SES</strong></caption>
<tr><td colspan="10" style="border-bottom: 1px solid black"></td></tr><tr><td style="text-align:left"></td
><td><strong>Group mean</strong></td><td colspan="8"><strong>Individual mathach</strong></td></tr>
<tr><td style="text-align:left"></td><td><strong>OLS</strong></td><td><strong>OLS</strong></td><td colspan
="7"><strong>linear</strong></td></tr>
<tr><td style="text-align:left"></td><td><strong></strong></td><td><strong></strong></td><td colspan="7"><
strong>mixed-effects</strong></td></tr>
<tr><td style="text-align:left"></td><td><strong>Between OLS</strong></td><td><strong>Pooled OLS</strong>
</td><td><strong>One-way-ANOVA</strong></td><td><strong>Random Intercept</strong></td><td><strong>Random C
oefficient</strong></td><td><strong>Intercept-as-Outcome</strong></td><td><strong>Contextual Effects</stro
ng></td><td><strong>Slope-as-Outcome</strong></td><td><strong>Slope-as-Outcome with Random Coeff</strong>
</td></tr>
<tr><td style="text-align:left"></td><td><strong>Model 1</strong></td><td><strong>Model 2</strong></td><td
><strong>Model 3</strong></td><td><strong>Model 4</strong></td><td><strong>Model 5</strong></td><td><stron
g>Model 6</strong></td><td><strong>Model 7</strong></td><td><strong>Model 8</strong></td><td><strong>Model
9</strong></td></tr>
<tr><td colspan="10" style="border-bottom: 1px solid black"></td></tr><tr><td style="text-align:left">SES
</td><td></td><td>2.95<sup>***</sup></td><td></td><td></td><td></td><td>2.39<sup>***</sup></td><td></td><t
d></td><td></td></tr>
<tr><td style="text-align:left"></td><td></td><td>(0.10)</td><td></td><td></td><td></td><td>(0.12)</td><td
></td><td></td><td></td></tr>
<tr><td style="text-align:left">Mean SES</td><td>5.39<sup>***</sup></td><td></td><td></td><td></td><td></t
d><td></td><td>5.87<sup>***</sup></td><td></td><td>5.24<sup>***</sup></td></tr>
<tr><td style="text-align:left"></td><td>(0.38)</td><td></td><td></td><td></td><td></td><td></td><td>(0.3
6)</td><td></td><td>(0.36)</td></tr>
<tr><td style="text-align:left">Group Mean Centered SES</td><td></td><td></td><td></td><td></td><td>2.19<s
up>***</sup></td><td></td><td>2.19<sup>***</sup></td><td>2.80<sup>***</sup></td><td>2.79<sup>***</sup></td
></tr>
<tr><td style="text-align:left"></td><td></td><td></td><td></td><td></td><td>(0.13)</td><td></td><td>(0.1
1)</td><td>(0.15)</td><td>(0.15)</td></tr>
<tr><td style="text-align:left">Grand Mean Centered SES</td><td></td><td></td><td></td><td>2.39<sup>***</s
up></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td style="text-align:left"></td><td></td><td></td><td></td><td>(0.11)</td><td></td><td></td><td></td>
<td></td><td></td></tr>
<tr><td style="text-align:left">Sector (1 = Catholic)</td><td>1.22<sup>***</sup></td><td>1.94<sup>***</sup
></td><td></td><td></td><td></td><td>2.54<sup>***</sup></td><td></td><td>2.81<sup>***</sup></td><td>1.25<s
up>***</sup></td></tr>
<tr><td style="text-align:left"></td><td>(0.32)</td><td>(0.15)</td><td></td><td></td><td></td><td>(0.34)</
td><td></td><td>(0.44)</td><td>(0.30)</td></tr>
<tr><td style="text-align:left">Group Mean Centered SES*Sector</td><td></td><td></td><td></td><td></td><td
></td><td></td><td></td><td>-1.34<sup>***</sup></td><td>-1.35<sup>***</sup></td></tr>
<tr><td style="text-align:left"></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td>
(0.23)</td><td>(0.23)</td></tr>
<tr><td style="text-align:left">Constant</td><td>12.12<sup>***</sup></td><td>11.79<sup>***</sup></td><td>1
2.64<sup>***</sup></td><td>12.66<sup>***</sup></td><td>12.64<sup>***</sup></td><td>11.47<sup>***</sup></td
><td>12.68<sup>***</sup></td><td>11.39<sup>***</sup></td><td>12.12<sup>***</sup></td></tr>
<tr><td style="text-align:left"></td><td>(0.20)</td><td>(0.11)</td><td>(0.24)</td><td>(0.19)</td><td>(0.2
4)</td><td>(0.23)</td><td>(0.15)</td><td>(0.29)</td><td>(0.20)</td></tr>
<tr><td style="text-align:left">Observations</td><td>160</td><td>7185</td><td>7185</td><td>7185</td><td>71
85</td><td>7185</td><td>7185</td><td>7185</td><td>7185</td></tr>
<tr><td style="text-align:left">R-squared</td><td>0.65</td><td>0.15</td><td></td><td></td><td></td><td></t
d><td></td><td></td><td></td></tr>
<tr><td style="text-align:left">Adj. R-squared</td><td>0.64</td><td>0.15</td><td></td><td></td><td></td><t
d></td><td></td><td></td><td></td></tr>
<tr><td style="text-align:left">Log Likelihood</td><td></td><td></td><td>-23558.40</td><td>-23322.58</td><
td>-23357.12</td><td>-23300.93</td><td>-23281.90</td><td>-23316.94</td><td>-23254.04</td></tr>
<tr><td style="text-align:left">AIC</td><td></td><td></td><td>47122.79</td><td>46653.17</td><td>46726.23</
td><td>46615.85</td><td>46573.81</td><td>46649.88</td><td>46526.08</td></tr>
<tr><td style="text-align:left">BIC</td><td></td><td></td><td>47143.43</td><td>46680.69</td><td>46767.51</
td><td>46664.01</td><td>46608.21</td><td>46704.92</td><td>46587.99</td></tr>
```

```
<tr><td colspan="10" style="border-bottom: 1px solid black"></td></tr><tr><td colspan="10" style="text-ali
gn:left"><sup></sup>p < .05; <sup></sup>p < .01; <sup></sup>p < .001</td></tr>
</table>
```

Model comparison over Mathach ~ SES

	Group mean OLS	OLS	Individual mathach linear mixed-effects						
	Between OLS	Pooled OLS	One-way-ANOVA	Random Intercept	Random Coefficient	Intercept-as-Outcome	Contextual Effects	Slope-as-Outcome	Slope-as-Outcome with Random Coeff
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9
SES		2.95^{***}				2.39^{***}			
		(0.10)				(0.12)			
Mean SES	5.39^{***}						5.87^{***}		5.24^{***}
	(0.38)						(0.36)		(0.36)
Group Mean Centered SES					2.19^{***}		2.19^{***}	2.78^{***}	2.79^{***}
					(0.13)		(0.11)	(0.14)	(0.15)
Grand Mean Centered SES				2.39^{***}					
				(0.11)					
Sector (1 = Catholic)	1.22^{***}	1.94^{***}				2.54^{***}		2.81^{***}	1.25^{***}
	(0.32)	(0.15)				(0.34)		(0.44)	(0.30)
Group Mean Centered SES*Sector								-1.35^{***}	-1.35^{***}
								(0.22)	(0.23)
Constant	12.12^{***}	11.79^{***}	12.64^{***}	12.66^{***}	12.64^{***}	11.47^{***}	12.68^{***}	11.39^{***}	12.12^{***}
	(0.20)	(0.11)	(0.24)	(0.19)	(0.24)	(0.23)	(0.15)	(0.29)	(0.20)
Observations	160	7185	7185	7185	7185	7185	7185	7185	7185
R-squared	0.65	0.15							
Adj. R-squared	0.64	0.15							
Log Likelihood			-23558.40	-23322.58	-23357.12	-23300.93	-23281.90	-23322.86	-23254.04
AIC			47122.79	46653.17	46726.23	46615.85	46573.81	46657.71	46526.08
BIC			47143.43	46680.69	46767.51	46664.01	46608.21	46698.99	46587.99

p < .05; p < .01; p < .001

Model Specification Summary (Cont’d)

::: {style="font-size: 0.55em;"}

Model comparison over Mathach ~ SES

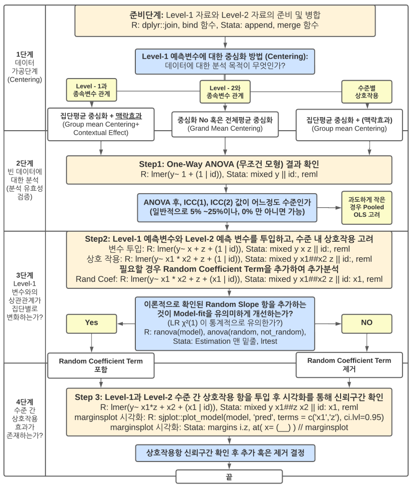
	OLS (dv = mathch)			HLM (dv = mathch)		
Predictors	*Estimates*	*std. Error*	*CI*	*Estimates*	*std. Error*	*CI*
ses	2.95^{***}	0.10	2.76 – 3.14			
ses groupmc				2.79^{***}	0.15	2.48 – 3.09
meanses				5.24^{***}	0.36	4.53 – 5.96
sector [Catholic]	1.94^{***}	0.15	1.64 – 2.23	1.25^{***}	0.30	0.66 – 1.85
ses groupmc × sector [Catholic]				-1.35^{***}	0.23	-1.81 – -0.89
(Intercept)	11.79^{***}	0.11	11.59 – 12.00	12.12^{***}	0.20	11.73 – 12.50
Random Effects						
σ^2				36.71		
τ_{00}				2.32 $_{id}$		
τ_{11}				0.25 $_{id.ses\_groupmc}$		
ρ_{01}				0.26 $_{id}$		
ICC				0.06		
N				160 $_{id}$		
Observations	7185			7185		
R^2 / R^2 adjusted	0.149 / 0.149			0.176 / 0.227		
AIC	46946.637			46532.575		

log-Likelihood	-23469.318	-23254.038

- *p<0.05 ** p<0.01 *** p<0.001*

:::

Module III: Sum-up



E.O.D.

참고: R vs. STATA (Summary & Fit Statistics)

```
Hierarchical Linear Model (HLM)
(also known as) Linear Mixed Model (LMM)
(also known as) Multilevel Linear Model (MLM)

Model Information:
Formula: mathach ~ ses_groupmc * sector + meanses + (ses_groupmc | id)
Level-1 Observations: N = 7185
Level-2 Groups/Clusters: id, 160

Model Fit:
AIC = 46526.077
BIC = 46587.994
R_(m)² = 0.17615  (Marginal R²: fixed effects)
R_(c)² = 0.22721  (Conditional R²: fixed + random effects)
Omega² = 0.23844  (= 1 - proportion of unexplained variance)
```

```
. mixed mathach c.ses_groupmc##ib0.sector meanses|| id: ses_groupmc, mle var cov(un)

Performing EM optimization ...

Performing gradient-based optimization:
Iteration 0:   log likelihood = -23254.609
Iteration 1:   log likelihood =  -23254.04
Iteration 2:   log likelihood = -23254.038

Computing standard errors ...

Mixed-effects ML regression                     Number of obs     =      7,185
Group variable: id                              Number of groups  =        160
                                                Obs per group:
                                                              min =         14
                                                              avg =       44.9
                                                              max =         67
                                                Wald chi2(4)      =     707.99
Log likelihood = -23254.038                     Prob > chi2       =     0.0000
. estat ic

Akaike's information criterion and Bayesian information criterion

-----------------------------------------------------------------------------
       Model |          N   ll(null)  ll(model)      df         AIC        BIC
-------------+---------------------------------------------------------------
           . |      7,185          .  -23254.04       9    46526.08   46587.99
-----------------------------------------------------------------------------
Note: BIC uses N = number of observations. See [R] BIC note.
```

참고: R vs. STATA (Fixed Effects)

```
Fixed Effects:
Unstandardized Coefficients (b or γ):
Outcome Variable: mathach
──────────────────────────────────────────────────────────────────────────────
                              b/γ    S.E.     t    df     p        [95% CI of b/γ]
──────────────────────────────────────────────────────────────────────────────
(Intercept)                 12.116 (0.197) 61.38 163.1 <.001 *** [11.726, 12.506]
ses_groupmc                  2.788 (0.155) 18.01 145.6 <.001 *** [ 2.482,  3.094]
sectorCatholic               1.253 (0.303)  4.13 152.5 <.001 *** [ 0.654,  1.852]
meanses                      5.244 (0.365) 14.38 154.2 <.001 *** [ 4.524,  5.964]
ses_groupmc:sectorCatholic  -1.348 (0.233) -5.78 154.9 <.001 *** [-1.809, -0.887]
──────────────────────────────────────────────────────────────────────────────
'df' is estimated by Satterthwaite approximation.

Standardized Coefficients (β):
Outcome Variable: mathach
──────────────────────────────────────────────────────────────────────────────
                               β     S.E.     t    df     p          [95% CI of β]
──────────────────────────────────────────────────────────────────────────────
ses_groupmc                  0.268 (0.015) 18.01 145.6 <.001 *** [ 0.238,  0.297]
sectorCatholic               0.091 (0.022)  4.13 152.5 <.001 *** [ 0.048,  0.135]
meanses                      0.315 (0.022) 14.38 154.2 <.001 *** [ 0.272,  0.359]
ses_groupmc:sectorCatholic  -0.086 (0.015) -5.78 154.9 <.001 *** [-0.115, -0.056]
──────────────────────────────────────────────────────────────────────────────
```

mathach	Coefficient	Std. err.	z	P>¦z¦	[95% conf.	interval]
ses_groupmc	2.787651	.1548387	18.00	0.000	2.484173	3.091129
1.sector	1.252655	.3032032	4.13	0.000	.658388	1.846923
sector#c.ses_groupmc						
1	-1.3481	.2333073	-5.78	0.000	-1.805374	-.8908265
meanses	5.244067	.3646407	14.38	0.000	4.529384	5.958749
_cons	12.11605	.1973894	61.38	0.000	11.72917	12.50292

참고: R vs. STATA (Random Effects)

ICC | τ_{00} | τ_{11} | $\tau_{01}(\rho_{01})$ | σ^2

```
Random Effects:
──────────────────────────────────────────────
 Cluster   K  Parameter    Variance     ICC
──────────────────────────────────────────────
 id      160 (Intercept)   2.31794   0.05940
              ses_groupmc  0.24589
 Residual                 36.70766
──────────────────────────────────────────────
```

Random-effects parameters	Estimate	Std. err.	[95% conf.	interval]
id: Unstructured				
var(ses_groupmc)	.2462764	.2248209	.0411507	1.473901
var(_cons)	2.318041	.3610305	1.708237	3.145532
cov(ses_groupmc,_cons)	.1924682	.2154886	-.2298816	.6148181
var(Residual)	36.70748	.6256972	35.50139	37.95454

```
LR test vs. linear model: chi2(3) = 217.17                Prob > chi2 = 0.0000
```

Random Intercept + Coefficient Test

Random Coefficient Test

```
> ranova(model11)
ANOVA-like table for random-effects: Single term deletions

Model:
mathach ~ ses_groupmc + sector + meanses + (ses_groupmc ¦ id) + ses_groupmc:sector
                               npar logLik   AIC    LRT Df Pr(>Chisq)
<none>                            9 -23254 46526
ses_groupmc in (ses_groupmc ¦ id)    7 -23255 46524 2.2248  2     0.3288
```

```
> anova(model10, model11)
Data: data_merged
Models:
model10: mathach ~ ses_groupmc * sector + meanses + (1 ¦ id)
model11: mathach ~ ses_groupmc * sector + meanses + (ses_groupmc ¦ id)
        npar   AIC   BIC logLik deviance  Chisq Df Pr(>Chisq)
model10    7 46524 46572 -23255    46510
model11    9 46526 46588 -23254    46508 2.2248  2     0.3288
```

```
Conditional intraclass correlation
──────────────────────────────────────────────────────────
          Level │      ICC   Std. err.     [95% conf. interval]
────────────────┼─────────────────────────────────────────
             id │ .0593981   .0087957      .0443271   .0791687
──────────────────────────────────────────────────────────
Note: ICC is conditi                                variates.
```

Random Coefficient Test

```
. lrtest final_rand final_nonrand, stats

Likelihood-ratio test
Assumption: final_nonrand nested within final_rand

 LR chi2(2) =   2.22
Prob > chi2 = 0.3288
```

Akaike's information criterion and Bayesian information criterion

Model	N	ll(null)	ll(model)	df	AIC	BIC
final_nonr~d	7,185	.	-23255.15	7	46524.3	46572.46
final_rand	7,185	.	-23254.04	9	46526.08	46587.99

참고: REML vs. MLE

- REML vs. MLE
  - REML이 Mixed Effect 모형 추정시 default 추정 방법
  - REML이나 MLE나 비슷한 회귀계수를 추정함
  - REML과 MLE의 경우 variance component 추정에서 차이가 존재
  - Level-2 unit들의 수가 적을때, MLE 분산에 대한 추정치가 REML보다 작게 계산되고, 결과적으로 좁은신뢰 구간과 biased된 유의성 검정으로 이어짐
- Likelihood Ratio test for nested models로 검정
  - 표본의 수가 작은 경우에 REML 사용을 통해 MLE를 보정, 표본의 수가 충분히 많을 경우 MLE
  - 만약 두 모형의 fixed effect들의 값이 동일하고, random effect의 값이 적다면 REML과 MLE는 혼용가능
  - 만약 두 모형의 fixed effect들의 값이 다르고, random effect의 값이 적다면 MLE를 사용해야 함

- 현실적으로는 수준 간 상호작용 같은 경우 MLE 사용 - estimation 속도

참고 (샘플1)

TABLE 3
Results of Hierarchical Linear Modeling Analyses of Expatriate Work Adjustment and Job Performance[a]

Variables	Work Adjustment Model 1	Work Adjustment Model 2	Work Adjustment Model 3	Job Performance Model 4	Job Performance Model 5	Job Performance Model 6
Level 1 main effects						
Age	.00 (.00)	.00 (.00)	.00 (.00)	−.01* (.00)	−.01* (.00)	−.01* (.00)
Marital status	−.07 (.08)	−.07 (.07)	−.07 (.07)	.20* (.08)	.19* (.08)	.20* (.08)
Prior Intl. experience	.00 (.01)	.00 (.01)	.00 (.01)	.01 (.01)	.01 (.01)	.01 (.01)
Assignment tenure in years	.02* (.01)	.02* (.01)	.02* (.01)	.03* (.01)	.03* (.01)	.03* (.01)
Language proficiency	.01 (.02)	.01 (.02)	.01 (.02)	.01 (.02)	.01 (.02)	.00 (.02)
Openness to experience	.38* (.06)	.31* (.07)	.30* (.06)	.04 (.07)	−.01 (.07)	−.05 (.07)
Emotional stability	.39* (.05)	.35* (.05)	.33* (.05)	−.01 (.05)	−.03 (.06)	−.09 (.06)
Job performance, 2006	.03 (.03)	.02 (.03)	.01 (.03)	.17* (.04)	.17* (.04)	.17* (.04)
Perceived support	.09* (.04)	.07 (.04)	.05 (.04)	.05 (.05)	.04 (.05)	.03 (.05)
Perceived cultural distance	−.06* (.03)	−.06* (.03)	−.06* (.03)	−.02 (.04)	−.02 (.03)	−.01 (.03)
Cross-cultural motivation		.21* (.05)	.27* (.06)		.12* (.05)	.09 (.05)
Work adjustment						.15* (.04)
Level 2 main effects						
Subsidiary support	−.09 (.16)	−.06 (.16)	−.04 (.17)	.27 (.15)	.28 (.15)	.28 (.15)
Cultural distance	.01 (.06)	.02 (.07)	.00 (.07)	.04 (.06)	.05 (.05)	.05 (.06)
Cross-level interactions						
Cross-cultural motivation × subsidiary support			−.66* (.23)			
Cross-cultural motivation × cultural distance			−.24* (.06)			
Pseudo R^2	.19	.20	.22	.09	.10	.11

[a] n = 556 expatriates (level 1) in 31 host countries/foreign subsidiaries (level 2). Unstandardized estimates (based on grand-mean centering) are reported, with standard errors in parentheses. Pseudo R^2 values estimate the amount of total variance (both level 1 and level 2) in the dependent variable captured by predictors in the model.
* $p < .05$
Two-tailed test.

prior research on overall support climate (McAllister & Bigley, 2002; Takeuchi et al., 2009). Both intermember reliability indexes (ICC1 = .06, ICC2 = .55, $F_{30,\ 555}$ = 2.20, $p < .05$) and interrater agreement (median $r_{wg(j)}$ = .95) provided support for aggregating individual support scores to the subsidiary level.

Level-2 변수의 투입에 관심

(Chen et al, 2010) AOM

참고 (샘플2)

Table 3 Results of Multilevel Analysis of the Effect of Monitoring and Sanctioning and Compliant (Noncompliant) Behavior of Leaders and Peers on Respondents' Willingness to Refuse Bribes

	Null Model Coeff.	Null Model SE	Model 1 Coeff.	Model 1 SE	Model 2 Coeff.	Model 2 SE	Model 3 Coeff.	Model 3 SE	Model 4 Coeff.	Model 4 SE
Fixed Part										
Constant	3.43***	.04	3.47***	.05	3.00***	.05	2.99***	.07	2.96***	.47
Level 1 Main design parameters										
Monitoring and sanctioning weak			−.26***	.04	−.25***	.04	−.31**	.11	−.43***	.11
Monitoring strong and sanctioning weak			−.13**	.04	−.13**	.04	−.10	.08	−.10	.08
Monitoring and sanctioning strong			.25***	.04	.25***	.04	.34**	.11	.23*	.11
Peer refused gift					.28***	.04	.27*	.11	.15	.11
Leader refused gift					.61***	.04	.53***	.11	.43***	.11
Peer refused gift * Leader refused gift					.08	.06	.33·	.18	.55***	.18
Level 1 Interaction parameters										
Monitoring and sanctioning weak* Peer refused gift							.18	.18	.40*	.18
Monitoring strong and sanctioning weak* Peer refused gift							−.12	.12	−.11	.12
Monitoring and sanctioning strong* Peer refused gift							−.03	.18	.17	.18
Monitoring and sanctioning weak* Leader refused gift							.15	.18	.36*	.18
Monitoring strong and sanctioning weak* Leader refused gift							.24*	.12	.23*	.12
Monitoring and sanctioning strong* Leader refused gift							−.04	.18	.15	.18
Monitoring and sanctioning weak* Peer and leader refused gift							−.44	.32	−.85**	.31
Monitoring strong and sanctioning weak* Peer and leader refused gift							−.37*	.16	−.37*	.17
Monitoring and sanctioning strong* Peer and leader refused gift							−.19	.32	−.58·	.31
Level 2 Respondent characteristics										
Trust in management									.18·	.10
Trust in peer									.21*	.09
Work relation with leader									.25***	.06
Work relation with peer									−.03	.07
Reward satisfaction									.03	.03
Job satisfaction									.02	.04
Knowledge of unethical cases in organization									.04	.05
Perception on scenarios presented									−.06	.06
Gender (ref: male)										
Female									.23*	.11
Education (ref: undergraduate)										
Graduate degree									−.07	.09
Government level (ref: central government)										
Local government (province and district)									.03	.12
Work experience in present position (ref: 0–12 months)										
> 12 months									−.14·	.08
Number of staff (ref: 0–50)										
51–100									−.03	.09
>100									.16·	.09
Position (ref: diplomat and other position)										
Position at local government									−.05	.15
Position at central government									.00	.15
Central at local government									−.00	.23
Control function									−.44	.34
Random Part										
Level 2: Respondent	.64	.05	.65	.05	.67	.05	.67	.05	.57	.04
Level 1: Vignette	1.15	.03	1.11	.03	.96	.02	.95	.02	.96	.02
−2*log likelihood:	14,684.2		14,542.0		13,953.4		13,923.2		13,442.0	
N Respondent	577		577		577		577		557	
N Vignette	4,602		4,602		4,602		4,602		4,452	

Note: · $p < .10$; * $p < .05$; ** $p < .01$; *** $p < .001$ (two tailed).
SCS: Senior civil servant.

Level-1 Interaction

L2 Intercept: τ_{00}

Deviance

L1 Intercept : σ^2

(Silitonga et al, 2019) PAR

참고 (샘플3)

Table 3. Multilevel Analysis of Certifier Sector and Respondent-Level Characteristic Relationships With Regulatory Approach Constructs

	Strictness of Regulatory Interpretation	Administer Sanctions	Flexibility	Provide Technical Assistance
Level 2: certifier attributes				
Nonprofit certifier (reference = for-profit)	−0.06 (0.08)	0.13 (0.09)	0.17 (0.11)	−0.08 (0.13)
Public certifier (reference = for-profit)	−0.13 (0.09)	0.17 (0.11)	−0.13 (0.13)	−0.37 (0.15)**
State market share	0.05 (0.15)	−0.31 (0.16)*	0.09 (0.20)	0.11 (0.21)
Number of clients (log transformed)	−0.04 (0.03)	0.08 (0.04)**	−0.11 (0.05)**	−0.11 (0.06)*
Level 1: respondent attributes				
State HHI	−0.02 (0.19)	0.31 (0.21)	−0.10 (0.26)	0.12 (0.29)
Profit motive	0.02 (0.02)	0.05 (0.02)**	0.01 (0.03)	−0.03 (0.03)
Organic movement motive	0.10 (0.02)***	0.03 (0.02)	0.07 (0.03)**	0.07 (0.03)**
Regulatory assessment	0.11 (0.03)***	0.03 (0.03)	0.07 (0.04)*	0.16 (0.04)***
Regulatory experience	0.02 (0.03)	0.07 (0.03)**	−0.02 (0.04)	−0.02 (0.04)
Predisposition towards laws	0.00 (0.03)	0.05 (0.03)	−0.04 (0.04)	−0.01 (0.04)
Operation size	0.03 (0.03)	0.08 (0.03)**	−0.06 (0.04)	−0.04 (0.04)
Pre-NOP certification	−0.04 (0.06)	0.08 (0.06)	−0.05 (0.08)	−0.12 (0.08)
Handling certification scope	0.07 (0.07)	0.00 (0.08)	0.09 (0.10)	−0.08 (0.10)
Constant	1.00 (0.25)***	−0.18 (0.30)	0.92 (0.38)**	1.31 (0.43)***
Percent variation explained by level 2 (certifier)	0.0%	1.9%	1.6%	2.8%
Log likelihood	−869	−911	−1,076	−1,112
N	784	766	761	777

Note: SEs in parentheses; level 2 (certifier) $N = 41$.
*Significant at $p < .10$; **significant at $p < .05$; ***significant at $p < .01$.

참고 (샘플4)

Table 2. Summary of hierarchical linear modelling results.

	Model 1			Model 2			Model 3			Model 4		
	β	*SE*	*t*	*β*	*SE*	*t*	*β*	*SE*	*t*	*β*	*SE*	*t*
Control variables												
(Level 1)												
Intercept	–.03	.05	–.54	–.03	.05	–.54	–.03	.05	–.55	–.02	.05	–.48
Gender	–.11**	.02	–4.26	–.08**	.02	–3.46	–.08**	.02	–3.45	–.08**	.02	–3.48
Age	.07	.06	1.29	.07	.05	1.33	.07	.05	1.32	.07	.05	1.37
Education	.02	.02	.82	.01	.02	.59	.01	.02	.58	.01	.02	.5
Work experience	.02	.06	.27	.01	.05	.24	.01	.05	.25	.01	.05	.22
Job rank	–.01	.03	–.18	–.03	.03	–1.08	–.03	.03	–1.06	–.03	.03	–1.11
(Level 2)												
Organization type	–.09	.05	–1.69	–.05	.05	–1.09	–.05	.05	–.89	–.05	.05	–1.02
Predictor variables												
(Level 1)												
Peer satisfaction				.46**	.02	21.79	.46**	.02	21.78	.45**	.02	21.42
(Level 2)												
Agency power							–.01	.05	–.17	.00	.05	–.03
Interaction effect												
Peer satisfaction × Agency power										.05**	.02	2.55
N		1781			1781			1781			1781	
R^2		.03			.23			.23			.24	
AIC		4957.05			4544.07			4549.73			4551.62	

* $p < .05$, ** $p < .01$.

Cross-level Interaction

Fit

(Kim and Eun, 2022) JAPP