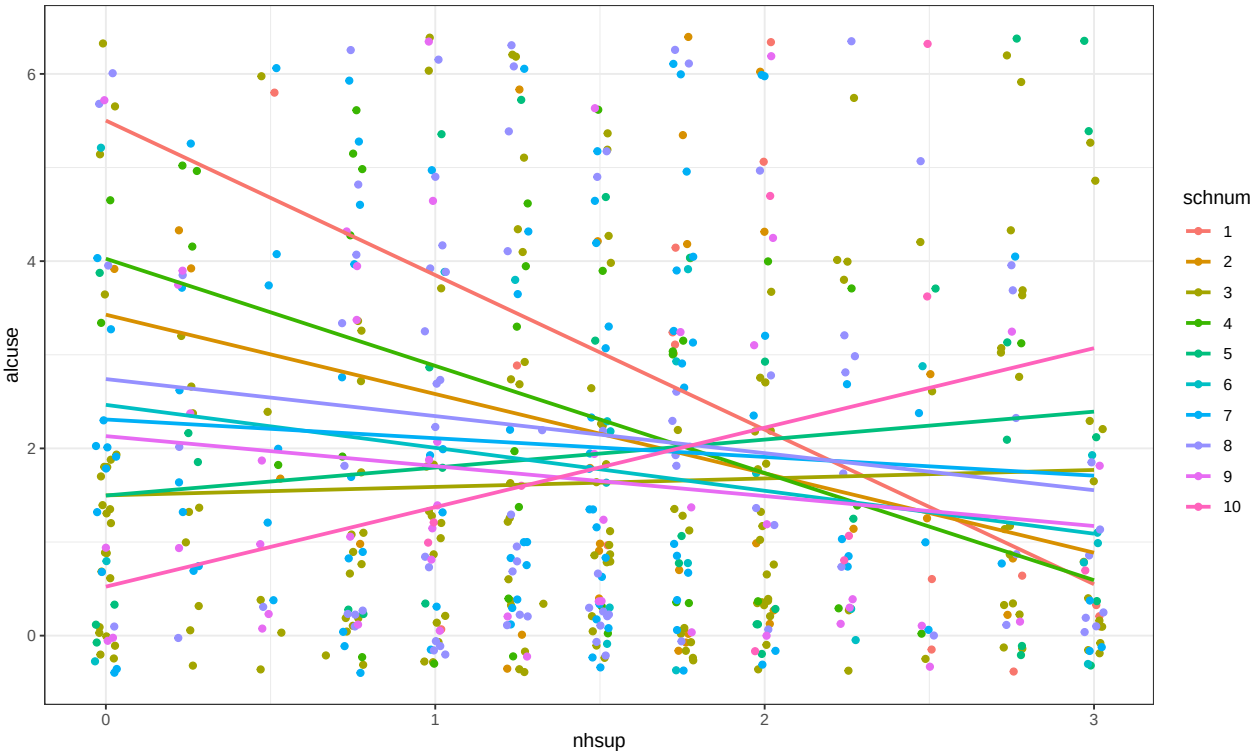
Homework 2

Adon Rosen

3/11/2021

Question 1

Part 1



Part 2

A table with the mean, standard deviation, standard error and the 95% confidence interval for the intercept and slopes can be found below

variable	N	value	sd	se	ci
Intercept	28	2.4519920	0.9455264	0.1786877	0.3666369
Slope	28	-0.4029244	0.4874196	0.0921136	0.1890016

Part 3

There appears to be considerable variation in both the intercept as well as the slope when relating the neighborhood support to alcuse within and between schools. This makes these data a good candidate for a mixed effects model given the fact that the schools represent a large normally distributed population and we have a subsample of this population, furthermore, every individual in the data set belongs to a single school.

Question 2

Part 1

The fitted model is:

$Y_{ij} = \beta_{0j} + r_{ij}$

Where:

Y_{ij} is the predicted value alcohol use for individual j from school i

$\beta_0 i$ is the school specific effect

r_{ij} is the individual specific random deviation

The school specific model is fitted as:

$\beta_0 j = \gamma_{00} + u_{0j}$

γ_{00} is the grand mean

u_{0j} is the random deviation of school j from the grand mean

The code to fit the model is: lme4::lmer(alcuse ~ (1|schnum), data=in.dat)

Part 2

The fixed effects for this model include the grand mean of alcohol use across all schools and the value is:

	x
(Intercept)	1.862258

This value suggests that the average alcohol use across all schools lays between 1-2 times in the past year, and 3-5 times in the last year, although considerably closer to the latter.

The random effects include the school specific alcohol use and the values are:

grpvar	term	grp	condval	condsd
schnum	(Intercept)	1	0.0882237	0.2511041
schnum	(Intercept)	2	0.0726617	0.2205687
schnum	(Intercept)	3	-0.1895951	0.1293442
schnum	(Intercept)	4	0.3182548	0.2148700
schnum	(Intercept)	5	0.0622674	0.2162532
schnum	(Intercept)	6	-0.0738560	0.2352512
schnum	(Intercept)	7	0.1388181	0.1577909
schnum	(Intercept)	8	0.1926131	0.1666177
schnum	(Intercept)	9	-0.0706494	0.2001052
schnum	(Intercept)	10	0.0644255	0.2628062
schnum	(Intercept)	11	0.0337492	0.2191016
schnum	(Intercept)	12	-0.2092231	0.1898809
schnum	(Intercept)	13	0.4128294	0.1453329
schnum	(Intercept)	14	-0.1131113	0.1140371
schnum	(Intercept)	15	-0.0591381	0.1600068
schnum	(Intercept)	16	-0.1288871	0.2205687
schnum	(Intercept)	17	0.0225103	0.2058782
schnum	(Intercept)	18	-0.2732011	0.1763513
schnum	(Intercept)	20	-0.0311178	0.1344079
schnum	(Intercept)	21	-0.6692047	0.1889425
schnum	(Intercept)	22	-0.1121690	0.1611503
schnum	(Intercept)	23	-0.0675793	0.2317358
schnum	(Intercept)	24	0.4699412	0.1136298
schnum	(Intercept)	25	-0.0828464	0.1278755
schnum	(Intercept)	26	-0.0252478	0.1541250
schnum	(Intercept)	27	0.0530532	0.1862087
schnum	(Intercept)	28	-0.0918654	0.2046836
schnum	(Intercept)	29	0.2683437	0.2058782

These random intercepts reflect the deviation from the grand mean within each of the schools. The values range from -0.67 to 0.47 with a varaiance of 0.09. indicating that some schools are on average more than half of a point lower then the grand mean, and other schools are on average about half of a point higher in alcohol use.

Part 3

The ICC of alcohol use of schools is: 0.02. This value reflects the amount of variation explained by group membership.

The design effect reflects the amount of correction that needs to be provided to estimate more accurate standard error estimates. It reflects the impact to the statistic estimation that is impacted by larger ICC values. That is, as ICC values increase, the standard error estimates are are reduced to a greater extent.

The design effect is estimated by:

$1 + (n - 1)\rho$ where:

n is the average group sample size

ρ is the intraclass correlation

The design effect for these data is: $1 + (n - 1) * \rho = 1 + (81.53571 - 1) * 0.02 = 2.85$

Given the magnitude of the design effect, and prior rule suggestions, this is adequate justification for the use of HLM as opposed to OLS.

Question 3

Part 1

The fitted model is:

$Y_{ij} = \beta_{0j} + \beta_{1j} * neighborhoodSupport + r_{ij}$

Where:

Y_{ij} is the predicted value alcohol use for individual j from school i

$\beta_0 i$ is the school specific effect

r_{ij} is the individual specific random deviation

The school specific model is fitted as:

$\beta_0 j = \gamma_{00} + u_{0j}$

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

γ_{00} is the grand mean

u_{0j} is the random deviation of school j from the grand mean

γ_{10} is the covariance between the slope and the intercept

u_{1j} is the random deviation from the mean slope

The code used to fit the model is:

lme4::lmer(alcuse~(nhsupZG|schnum),data=in.dat)

Part 2

The variance of the school intercepts is: 0.09

Part 3

The variance of the slopes by neighborhood support across all schools is: 0.06

Part 4

The covariation of the intercepts and slopes is: -0.03. This indicates a negative relationship between the average alcohol use within a school and the effect that neighborhood support has within a school. So as the mean neighborhood alcohol use increases the impact of neighborhood support decreases.

Part 5

The proportion of the variance explained by the random effects is: 0.03. About 3% of the total variance observed in the outcome can be explained due to the nested nature of the data, as well as differential responses to neighborhood support based school membership.

Question 4

Part 1

The fitted first level model is: $Y_{ij} = \beta_0 + \beta_{1j}(neighborhoodSupport) + r_{ij}$

Where

Y_{ij} is equal to the predicted alcohol use for an individual

β_{0j} is the mean alcohol use for school j

β_{1j} is the impact of dropout rates for school j

r_{ij} is the random deviation for individual i from school j

The second level models includes:

$\beta_{0j} = \gamma_{00} + \gamma_{01} * dropout + u_{0j}$

$\beta_{1j} = \gamma_{10} + \gamma_{11} * dropout + u_{1j}$

where

γ_{00} is the mean intercept across all schools

γ_{01}mean difference

u_{0j} is a random deviation from the mean intercept

u_{1j} is a random deviation from the mean slope

The code to train the model is: lme4::lmer(alcuse~dropoutZ*nhsupZG+(nhsupZG|schnum),data=in.dat)

Part 2

The fixed effects include:

	x
(Intercept)	1.8571799
dropoutZ	-0.0338379
nhsupZG	-0.3569134
dropoutZ:nhsupZG	0.0322948

The intercept value indicates the average estimated alcohol use is 1.78 across all schools and individuals. The dropout coefficient is being used to model the variance in the random slope and intercept included in the first level model and is of a very small magnitude (-0.03) which reflects it explains a small amount of the variation in the intercepts and slopes from the first level models. The main effect of neighborhood support reflects average main effect that neighborhood support has when predicting alcohol use across all schools, it suggests that across all schools as neighborhood support increases alcohol use decreasess at a relatively strong rate (-0.37).

The random effects include:

grp	(Intercept)	nhsupZG
1	0.0935499	-0.1433631
2	0.0775811	-0.1033748
3	-0.2011767	0.3174524
4	0.3127050	-0.2192539
5	-0.0618180	0.2736456
6	-0.0942335	0.0167345
7	0.0990805	0.0746933
8	0.1180446	0.0065529
9	-0.0953866	0.0454333
10	0.0008263	0.0774387
11	0.0310767	0.0125587
12	-0.0899377	-0.1277463
13	0.3682642	-0.2879139
14	-0.0299261	0.0943847
15	-0.0990980	0.0835296
16	-0.1512748	-0.0303738
17	0.0252516	-0.0614787
18	-0.2827002	0.0350462
20	-0.0575346	-0.0502643
21	-0.6373211	0.0900496
22	-0.0984181	0.4168450
23	-0.0445881	-0.0244363
24	0.3761119	0.0156419
25	0.0702441	-0.0499092
26	-0.0051455	-0.1790611
27	0.1715589	-0.1200911
28	-0.1278633	-0.0189557
29	0.3321275	-0.1437841

The total variance explained by the fixed effects model is: 0.104 which suggests the dropout rate does not explain a noticeable portion of the total variation in alcohol use. This is also reflected in the magnitude of the standardized effect size of: -0.03.

The random effects variance is 0.119 and explains and it explains 0.03 percent of the variance, across the entire model.

Question 5

The fixed effects include:

	x
(Intercept)	2.5146187
nhsup	-0.3714392
gender	-0.2455784

The intercept reflects the grand mean of alcohol use across all schools. The neighborhood support reflects the mean covariance that the neighborhood support has across all schools. It suggests a similar effect as observed in question 4 where increased neighborhood support relates to lower alcohol use. Finally, the last fixed effect is that of gender, where it appears a gender value equal to 1 decreases alcohol use by .25 on average across all schools.

The random effects include:

grp	(Intercept)	nhsup	gender
1	0.3986111	-0.1415363	-0.0601191
2	0.3979928	-0.1252798	-0.2083692
3	-0.7445951	0.3388213	0.1169478
4	0.7398378	-0.2735991	-0.0972891
5	-0.4303165	0.2554103	0.1400152
6	0.0898364	-0.0375143	-0.0653900
7	0.0873879	0.0368430	-0.1319688
8	0.4076801	-0.0415699	-0.2330441
9	-0.0102277	0.0211155	-0.1857126
10	-0.0703988	0.0907919	-0.0532409
11	-0.0302679	0.0232068	0.0970036
12	-0.3848578	0.0076330	0.2962332
13	0.7375975	-0.3501371	0.1291937
14	-0.5591872	0.1936016	0.2861869
15	-0.0686275	0.0571812	-0.1114279
16	0.0029320	-0.0860430	0.1291397
17	0.2678369	-0.0384401	-0.3856133
18	-0.3734215	0.0338896	0.1624658
20	-0.0757811	-0.0502211	0.1082154
21	-0.4191285	0.0150325	-0.1900964
22	-0.7864951	0.4772978	-0.0993215
23	0.0996689	-0.0271964	-0.2183422
24	0.1697726	-0.0668725	0.6437718
25	-0.0983223	0.0773199	-0.2274435
26	0.2349824	-0.1834249	0.0031189
27	-0.1138430	-0.0196897	0.3240248
28	-0.0009148	-0.0624324	0.1086502
29	0.5322486	-0.1241877	-0.2775883

The variance-covariance matrix of the random effects is:

	(Intercept)	nhsup	gender
(Intercept)	0.3024011	-0.1160861	-0.0715688
nhsup	-0.1160861	0.0623397	0.0047868
gender	-0.0715688	0.0047868	0.1337283

This indicates that the direction of relationship between the mean school alcohol use with both gender and neighborhood support is negative such that as the mean alcohol use increases the strength of the slope for both the neighborhood support and gender effect decreases. Interestingly, there is little relationship between the gender and neighborhood support slopes.

Overall, the variance of the random effects is: 0.15 and it explains 0.04 percent of the variance, across the entire model.