Durham University
Research Methods Centre

MATH43515: Multilevel Modelling

Lecture 5: Advanced multilevel models

Module Convenor/Tutor:

Andy Golightly

Outline (Lecture 5)

- Assessing significances of fixed and random effects
- Deviance (Likelihood ratio) tests
- Models with more than one predictor variable
- Model simplification strategies
- Diagnostics
- 3-level models: Variance partition coefficients and ICC
- Fitting and interpreting 3-level-models

Assessing significances: Fixed effects

We still work in the framework of the random intercept and slope model from Slide 4 of Lecture 4, in the context of the popularity data (fitted on Slide 8,9)

For fixed effects, this is just analogous to the linear model (ordinary least squares). We have the two test problems

$$H_0: a = 0 \quad \text{vs} \quad H_1: a \neq 0 \qquad \text{and} \qquad H_0: b = 0 \quad \text{vs} \quad H_1: b \neq 0$$

The t-values are given by |Estimate/Std.Error|, in this case 12.12 and 19.36, respectively.

The p-values give the probabilities of "obtaining by chance an even more extreme value of t".

Both p-values are ≈ 0, so in both cases H_0 is clearly rejected, and we conclude that both fixed effects are needed in the model.

Assessing significances: Random effects

For the random effects, we are interested in testing

$$H_0\colon \sigma_u^2 = 0 \ \text{ versus } \ H_1\colon \sigma_u^2 \neq 0$$

and

$$H_0\colon \sigma_v^2 = 0 \ \text{ versus } \ H_1\colon \sigma_v^2 \neq 0$$

One could do just do the same as before; i.e. |Estimate/Std.Error|. However,

- the standard errors of σ_u^2 and σ_v^2 are not available from the model output (the values given at Std.Dev. are just σ_u and σ_v!);
- even if we had them available, the resulting test would be inaccurate since the sampling distribution of the variances is skewed.

Assessing significances: Random effects (cont'd)

Instead, for random effects, one should carry out chi-squared tests based on deviances.

This is achieved by fitting the model with and without the random effect term, and then comparing the difference in deviances with a χ^2 distribution with degrees of freedom given by the difference in parameters.

Note: Deviance is defined as $D = -2\log L + c$, with c being a constant which however cancels out when taking differences (and which in lmer is set to 0). So, when having two models M_0 and M_1, where M_0 is nested in M_1, then $L_0 < L_1$ and $D_0 > D_1$ with test statistic

$$D_0 - D_1 = -2\log L_0 + 2\log L_1 = -2\log\frac{L_0}{L_1}$$

so that this is equivalent to a likelihood ratio test.

Assessing significances: Random effects (cont'd)

Beginning with a test for the random slope,

$$H_0\colon \sigma_v^2 = 0 \quad \text{versus} \quad H_1\colon \sigma_v^2 \neq 0$$

```
> model0 <- lmer(formula = popularity ~ 1+
extraversion +(1|class), data    = pop.data)
> deviance(model0)
[1] 5832.639
> model1 <- lmer(formula = popularity ~ 1+
extraversion +(1+ extraversion|class), data    =
pop.data)
> deviance(model1)
[1] 5779.395
> deviance(model0)-deviance(model1)
[1] 53.24409
> qchisq(0.95,2)
[1] 5.991465
```

The deviance difference is given by

$$D_0 - D_1 = 53.244\,.$$

At the 5% level of significance, this is compared to the 95% quantile of the χ_2^2 distribution, which is 5.99. So, H_0 is clearly rejected, and we decide that the random slope is necessary.

Note: The deviance() function gives a warning message of being deprecated. The reason for this is that it could be mistakenly used to test for significance of *fixed* effects. An alternative to using deviance(model0) is

-2*summary(model0)$logLik

(https://github.com/lme4/lme4/issues/*211)*

Durham University

Deviance chi-squared test via ranova()

```
> ranova(model1)
ANOVA-like table for random-effects: Single term deletions
Model:
  popularity ~ extraversion + (1 + extraversion | class)
                                              npar  logLik    AIC
   <none>                                        6  -2889.7  5791.4
   extraversion in (1 + extraversion | class)    4  -2916.3  5840.6

                                              LRT     Df   Pr(>Chisq)
  <none>
   extraversion in (1 + extraversion | class) 53.244  2   2.743e-12 ***
```

Note:
$2 \times 2889.7 = 5779.4$
$2 \times 2916.3 = 5832.6$

Assessing significances: Random effects (cont'd)

We *could* now do the same for the random intercepts

$$H_0\colon \sigma_u^2 = 0 \quad \text{versus} \quad H_1\colon \sigma_u^2 \neq 0$$

However,

- We know from the model output (Slide 9 of Lec 4) that the variance of the random intercept is much larger than that of the slope;
- It would be considered unusual to fit a model *with* a random slope but *without* a random intercept.

So, given the previous result, the data analyst would usually refrain from doing this, unless there is a very strong reason (for instance, an equal baseline response for all groups in the study).

Models with more than one predictor variable

Include gender and teacher experience into model for the popularity data:

```
> model2 <-
      lmer(formula =
      popularity ~ 1 + gender + extraversion + experience
      + (1 + gender + extraversion | class),
      data    = pop.data)
```

Notes:

- A random slope for experience (class level variable) cannot (meaningfully) be included.
- We are not applying any centering here for simplicity.…

Models with more than one predictor variable (cont’d)

```
> summary(model 2)
Random effects:
 Groups   Name        Variance  Std.Dev. Corr
 class    (Intercept) 1.342110  1.15849
          gender      0.002404  0.04903  -0.39
          extraversion 0.034738 0.18638  -0.88 -0.09
 Residual             0.551437  0.74259
Number of obs: 2000, groups:  class, 100

Fixed effects:
              Estimate   Std. Error  df         t value  Pr(>|t|)
(Intercept)   7.586e-01  1.973e-01   1.810e+02  3.845    0.000167 ***
gender        1.251e+00  3.694e-02   9.861e+02  33.859   < 2e-16 ***
extraversion  4.529e-01  2.464e-02   9.621e+01  18.376   < 2e-16 ***
experience    8.951e-02  8.618e-03   1.013e+02  10.387   < 2e-16 ***
```

Durham University

Simplifying the model

We see now a very small variance for the random effect of the gender term (0.002404)

```
> model3 <- lmer(formula = popularity ~ 1 + gender + extraversion + experience + (1 + extraversion | class),
               data    = pop.data)
> deviance(model3)-deviance(model2)
[1]  1.513252
> qchisq(0.95,3)
[1] 7.814728
```

So, at this occasion we do remove the random effect for gender (and decide for model3)

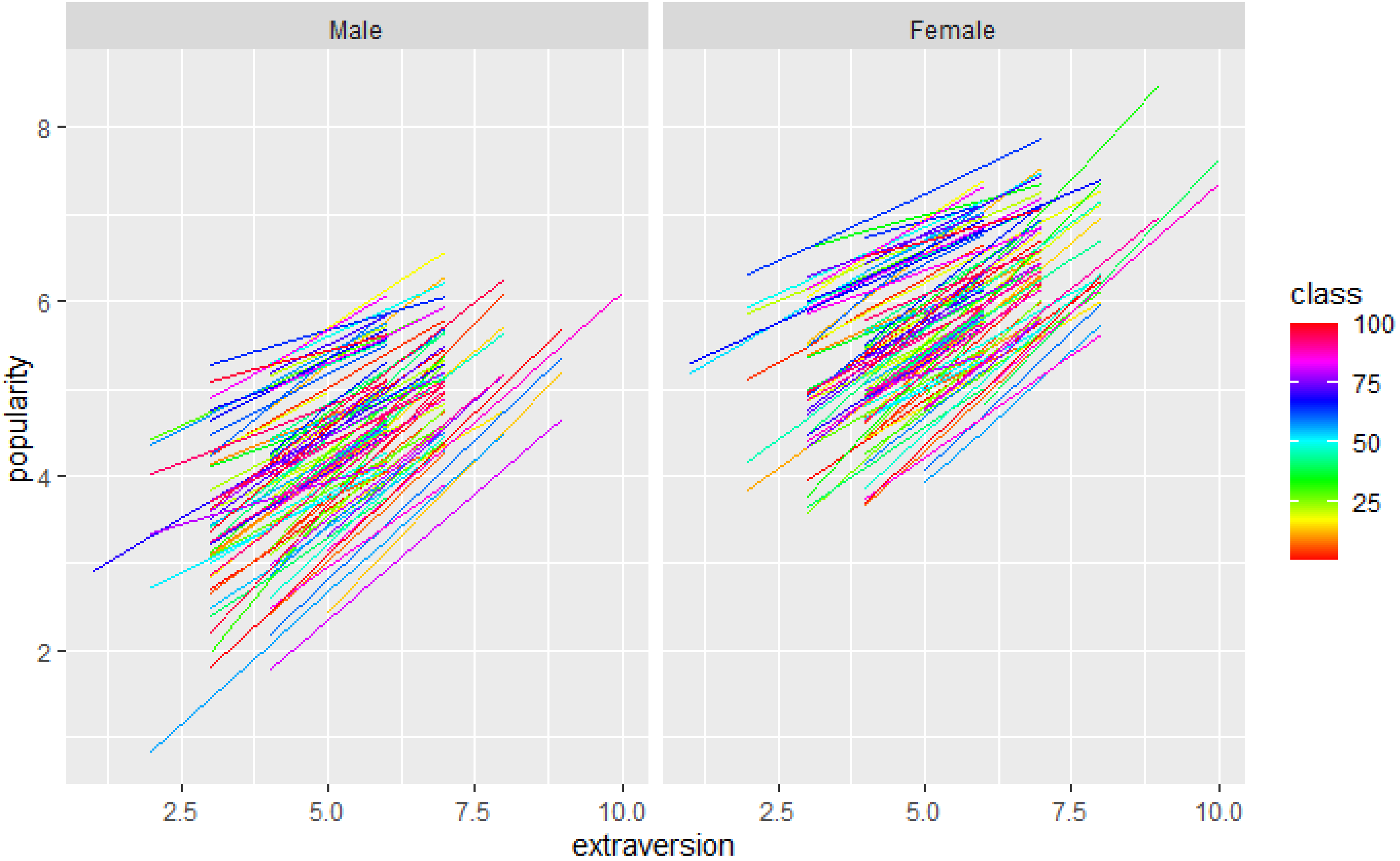
Final fitted model summaries

```
> summary(model3)
Random effects:
  Groups    Name         Variance  Std.Dev.  Corr
  class     (Intercept)  1.30299   1.1415
            extraversion 0.03455   0.1859    -0.89
  Residual               0.55209   0.7430
  Number of obs: 2000, groups:  class, 100

Fixed effects:
              Estimate   Std. Error  df         t value   Pr(>|t|)
  (Intercept)  7.361e-01  1.966e-01  1.821e+02  3.745     0.000242 ***
  gender       1.252e+00  3.657e-02  1.913e+03  34.240    < 2e-16 ***
  extraversion 4.526e-01  2.461e-02  9.754e+01  18.389    < 2e-16 ***
  experience   9.098e-02  8.685e-03  1.017e+02  10.475    < 2e-16 ***
```

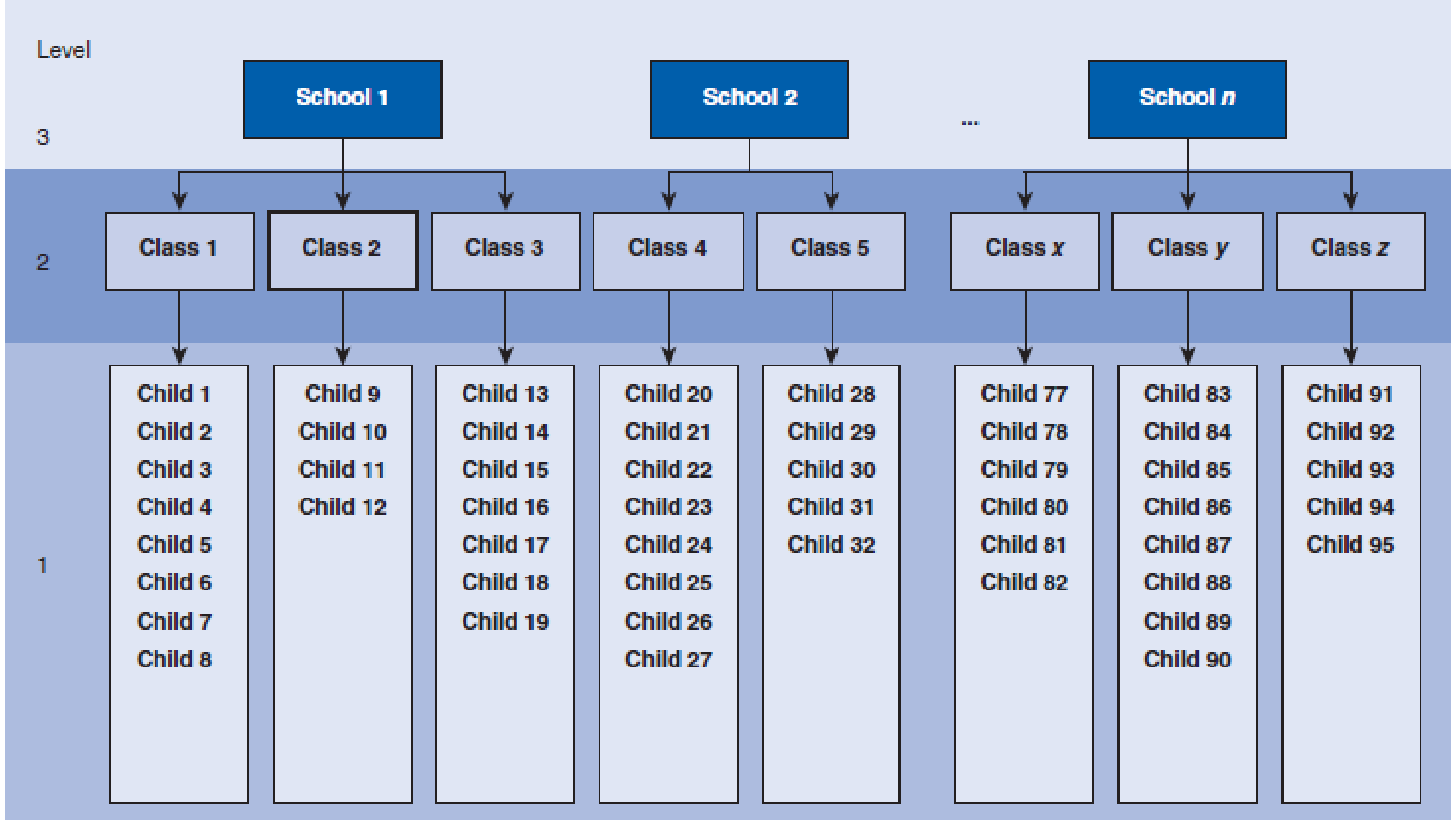
```
> names(summary(model3))
 [1] "methTitle"  "objClass"    "devcomp"   "isLmer"
 [5] "useScale"   "logLik"      "family"    "link"
 [9] "ngrps"      "coefficients" "sigma"    "vcov"
[13] "varcor"     "AICtab"      "call"      "residuals"
[17] "fitMsgs"    "optinfo"

> summary(model3)$varcor
  Groups   Name         Std.Dev.  Corr
  class    (Intercept)  1.14149
           extraversion 0.18587   -0.885
 Residual               0.74303

> round(summary(model3)$coef, digits=4)
              Estimate  Std. Error  df         t value   Pr(>|t|)
 (Intercept)   0.7361   0.1966      182.0991   3.7446    2e-04
 gender        1.2523   0.0366      1913.0946  34.2396   0e+00
extraversion   0.4526   0.0246      97.5389    18.3888   0e+00
experience     0.0910   0.0087      101.6558   10.4750   0e+00

> round(sqrt(diag(summary(model3)$vcov)), digits=4)
[1] 0.1966  0.0366  0.0246  0.0087
```

Durham University

Fitted models by gender



Three level data structures



Durham University

Three level empty model

$$y_{ijk} = \gamma_0 + u_{jk} + v_k + e_{ijk}$$

Terms	Description	
y_{ijk}	is the outcome variable for i^{th} person in j^{th} class and k^{th} school	
γ_0	is the overall Intercept	
u_{jk}	is the random effect for classroom	u_{jk} ~ N(0, σ_u^2)
v_k	is the random effect for school	v_k ~ N(0, σ_v^2)
e_{ijk}	student level residual error term	e_{ijk} ~ N(0, σ_e^2)
The school, classroom effects and the student level residual errors are assumed independent and normal distributed with zero means and constant variances.		

Note: We are operating in the module a slight notational simplification when compared to some literature as for instance Hox, who denotes $\gamma_0 = \gamma_{000}, u_{jk} = u_{0jk}, v_k = v_{00k}$

Variance partitions

Total variance is partitioned into three components:

- σ_v^2 variance component for level k (school k)
- σ_u^2 variance component for level j (class j)
- σ_e^2 individual variance, based on individual departures from group means (σ_e^2).

Variance partition coefficients (VPCs) report the proportion of the observed response variation that lies at each level of the model hierarchy.

The school level VPC is calculated as $\frac{\sigma_v^2}{\sigma_v^2+\sigma_u^2+\sigma_e^2}$

The classroom level VPC is calculated as $\frac{\sigma_u^2}{\sigma_v^2+\sigma_u^2+\sigma_e^2}$

The individual level VPC is calculated as $\frac{\sigma_e^2}{\sigma_v^2+\sigma_u^2+\sigma_e^2}$

Durham University

Three level model ICCs

Intraclass correlation coefficients (ICCs) measure the model implied correlation (i.e. similarity or homogeneity) of the observed responses within a given cluster

The kth school level ICC is calculated as the correlation between two individuals i and i' within the same school k, but different classrooms j and j'

$$\frac{\sigma_v^2}{\sigma_v^2 + \sigma_u^2 + \sigma_e^2}$$

The classroom level ICC is calculated as the correlation between two students i and i' within the same classroom j and therefore the same school k

$$\frac{\sigma_v^2 + \sigma_u^2}{\sigma_v^2 + \sigma_u^2 + \sigma_e^2}$$

We see that the classroom level ICC does not coincide with the classroom level VPC.

Durham University

Case study

Variables	Description
Math	Maths score
ActiveTime	Active time
ClassSize	Class size
Classroom	Class identifier
School	School identifier
StudentID	Student identifier

Fitting the empty three-level model

```
> Model.0 <- lmer(Math ~ 1+(1|School)+(1|School:Classroom),
                  data=Sim3level)
> summary(Model.0)
```

Here we have

- a random effect for schools: (1|School)
- a random effect for classes within schools: (1|School:Classroom)

Random intercept model estimates

	Empty model	
	Beta	SE
Intercept	44.42	5.67
Variances:		
Residual	37.22	
Class	44.93	
School	91.87	
VPC (class)	0.26	
VPC (school)	0.53	Variability between schools is more than between classes.
ICC (class)	0.79	Very large!!
ICC (school)	0.53	This *is* the same as VPC (school)!

Durham University

Three level model

Lets consider two important explanatory variables in the analysis.

$$Math_{ijk} = a + b_1 ActiveTime_{ijk} + b_2 Class_{jk} + u_{jk} + v_k + e_{ijk}$$

Here *ActiveTime* is at individual level (level 1)

The variable *Class* (class size) is at class level (level 2)

```
Model.1 <- lmer(Math ~ ActiveTime+ClassSize
        +(1|School)
        +(1|School:Classroom),
        data=Sim3level)
```

or

```
Model.1 <- lmer(Math ~ ActiveTime+ClassSize
        +(1|School/Classroom),
        data=Sim3level)
```

Durham University

Random intercept model estimates from an analysis with Math score as an outcome

	Empty model				Model with explanatory variables			
	Fixed effects	SE	T-value	P-value	Fixed effects	SE	T-value	P-value
Intercept	44.42	5.67	7.82	0.00	41.18	11.91	3.46	0.00
ActiveTime					14.95	0.61	24.64	0.00
Class size					-0.21	0.56	-0.37	0.71
Variances:								
Residual	37.22				17.54			
Class	44.93				46.55			
School	91.87				83.98			
Deviance	3781.8				3373.6			

Durham University

LR test

```
> anova( Model.0,  Model.1)
> Data: Sim3level
Models:
Model.0: Math ~ 1 + (1 | School) + (1 | School:Classroom)
Model.1: Math ~ ActiveTime + ClassSize + (1 | School/Classroom)
        npar    AIC    BIC  logLik deviance  Chisq Df Pr(>Chisq)
Model.0    4 3789.8 3807.2 -1890.9   3781.8
Model.1    6 3385.6 3411.7 -1686.8   3373.6 408.21  2  < 2.2e-16 ***
```

Results from intercept model analysis

- Parameter estimate for Active time is much larger than the corresponding standard error, and calculation of the t-test shows that it is significant at $p < 0.005$.
- Reduction in deviance indicates our model has improved by including both active time and class size. Evidence to suggest that class size isn't needed. Explore further in the lab.
- Reduction in residual (within class) variance indicates that individual level variable Active time explains some variance at the pupil level.

Thank you!!!!!

Durham University
Research Methods Centre

andrew.golightly@durham.ac.uk