Level-1 Error Structure

2017-11-07

Preparation

We will use the data in the *rat_weight_2.csv* file. These data are from Crowder & Hand (1990) and

> …describe data on the body weights of rats measured over 64 days. These data also appear in Table 2.4 of Crowder and Hand (1990). The body weights of the rats (in grams) are measured on day 1 and every seven days thereafter until day 64, with an extra measurement on day 44. The experiment started several weeks before "day 1." There are three groups of rats, each on a different diet.

```r
# Load libraries
library(tidyverse)
library(corrr)
library(lme4)
library(stringr)
library(AICcmodavg)


# This is to disable scientific notation
options(scipen = 99)

# Read in wide data
vocab = read_csv("~/Dropbox/epsy-8282/data/vocab.csv")

# Create long data
vocab_long = vocab %>%
  gather(time, score, t8:t11) %>%
  mutate(
    grade = as.integer(str_replace(time, pattern = "t", replacement = "")) - 8
    ) %>%
  arrange(id, grade)

head(vocab_long)
```

```
## # A tibble: 6 x 5
##      id female  time score grade
##   <int>  <int> <chr> <dbl> <dbl>
## 1     1      0    t8  1.75     0
## 2     1      0    t9  2.60     1
## 3     1      0   t10  3.76     2
## 4     1      0   t11  3.68     3
## 5     2      0    t8  0.90     0
## 6     2      0    t9  2.47     1
```

For this set of notes we will adopt a linear fixed-effect of grade. We will also adopt random-effects of both intercept and grade

$$\text{Score}_{ij} = \beta_0 + \beta_1(\text{Grade}_{ij}) + b_{0i} + b_{1i}(\text{Grade}_{ij}) + \epsilon_{ij}$$

$$\epsilon_{ij} \sim \mathcal{N}\left(0, \sigma_{\epsilon}^2\right)$$

$$\begin{bmatrix} b_{0i} \\ b_{1i} \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \left(\begin{bmatrix} \sigma_0^2 & \sigma_{01} \\ \sigma_{10} & \sigma_1^2 \end{bmatrix} \right) \right)$$

Fitting the Model

```
lmer.1 = lmer(score ~ 1 + grade + (1 + grade | id), data = vocab_long, REML = FALSE)
print(summary(lmer.1), show.resids = FALSE, correlation = FALSE)
```

```
## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula: score ~ 1 + grade + (1 + grade | id)
##    Data: vocab_long
##
##      AIC      BIC   logLik deviance df.resid
##    884.7    905.9   -436.3    872.7      250
##
## Random effects:
##  Groups   Name        Variance   Std.Dev. Corr
##  id       (Intercept) 3.13762586 1.771334
##           grade       0.00007514 0.008668 1.00
##  Residual             0.89633196 0.946748
## Number of obs: 256, groups:  id, 64
##
## Fixed effects:
##             Estimate Std. Error t value
## (Intercept)  1.41330    0.24255   5.827
## grade        0.74666    0.05294  14.105
```

When you fit the mixed-effects model using the `lmer()` function, one of the model assumptions is that the level-1 residuals has an independent error structure. Recall that the independent error structure implies a diagonal matrix with the SAME error variances on the diagonal, and all covariances are FIXED to zero. In a previous set of notes (matrix formulation) we defined this structure for each subject i as,

$$\mathbf{W_i} = \sigma_{\epsilon}^2 \mathbf{I_i}$$

where σ_{ϵ}^2 is the variance of the level-1 errors, and $\mathbf{I_i}$ is a $n_i \times n_i$ identity matrix. For our student with four time points,

$$\mathbf{W_i} = \begin{bmatrix} \sigma_{\epsilon}^2 & 0 & 0 & 0 \\ 0 & \sigma_{\epsilon}^2 & 0 & 0 \\ 0 & 0 & \sigma_{\epsilon}^2 & 0 \\ 0 & 0 & 0 & \sigma_{\epsilon}^2 \end{bmatrix}$$

To evaluate whether the empirical data correspond to this error structure, we will compute the correlations between the models' residuals at each time point. If the independence assumption is satisfied, these correlations should be zero (or very near zero).

2

```
out = broom::augment(lmer.1)
head(out)
```

```
##   score grade id  .fitted       .resid   .fixed       .mu .offset .sqrtXwt
## 1  1.75     0  1 1.797158 -0.04715805 1.413297 1.797158       0        1
## 2  2.60     1  1 2.545693  0.05430726 2.159953 2.545693       0        1
## 3  3.76     2  1 3.294227  0.46577257 2.906609 3.294227       0        1
## 4  3.68     3  1 4.042762 -0.36276212 3.653266 4.042762       0        1
## 5  0.90     0  2 1.206275 -0.30627544 1.413297 1.206275       0        1
## 6  2.47     1  2 1.951919  0.51808138 2.159953 1.951919       0        1
##   .sqrtrwt .weights       .wtres
## 1        1        1 -0.04715805
## 2        1        1  0.05430726
## 3        1        1  0.46577257
## 4        1        1 -0.36276212
## 5        1        1 -0.30627544
## 6        1        1  0.51808138
```

```
out %>%
  select(id, grade, .resid) %>%  # Only work with the id, grade and .resid columns
  spread(grade, .resid) %>%      # Convert these columns to wide data
  select(-id) %>%                # Drop the Rat column
  correlate() %>%                # Compute correlation matrix
  fashion(decimals = 2)
```

```
##   rowname   X0   X1   X2   X3
## 1       0      -.26 -.08 -.43
## 2       1 -.26      -.39 -.38
## 3       2 -.08 -.39      -.26
## 4       3 -.43 -.38 -.26
```

As you see from the correlation matrix, there are some high correlations among the level–1 residuals.

Fitting Models with Alternative Level–1 Error Structures

Although the independence assumption is very commonly used in practice, it is possible to allow a more flexible error structure. There are many different error structures you can impose on the error variance-covariance matrix.

Unfortunately, the lmer() function is only able to fit the independent error structure. To fit other error structures, we will use the lme() function from the **nlme** package. The syntax for lme() is very similar to that for lmer(). WIth the lme() function we specify the random-effects using a formula in the random= argument. Here we will again fit the same mixed-effects model as we did with lmer() using the lme() function.

```r
library(nlme)

# Fit model with independent level-1 error structure
lme.independence = lme(score ~ 1 + grade, random = ~ 1 + grade | id,
                       data = vocab_long)

summary(lme.independence)
```

```
## Linear mixed-effects model fit by REML
##   Data: vocab_long
##        AIC      BIC    logLik
##   889.8282 911.0522 -438.9141
##
## Random effects:
##  Formula: ~1 + grade | id
##  Structure: General positive-definite, Log-Cholesky parametrization
##             StdDev        Corr
## (Intercept) 1.79909581752 (Intr)
## grade       0.00007674295 0
## Residual    0.94929009247
##
## Fixed effects: score ~ 1 + grade
##                 Value  Std.Error  DF   t-value p-value
## (Intercept) 1.4132969 0.24582616 191  5.749172       0
## grade       0.7466562 0.05306693 191 14.070086       0
##  Correlation:
##       (Intr)
## grade -0.324
##
## Standardized Within-Group Residuals:
##         Min          Q1         Med          Q3         Max
## -2.33786417 -0.65487369 -0.02911057  0.64249099  2.64835556
##
## Number of Observations: 256
## Number of Groups: 64
```

By default, the lme() function also fits the independent error structure. So if you compare the fixed effects and random effects using the functions lmer() and lme(), you will notice they are very similar.

Small differences are usual, since the lmer() and lme() may be using different estimation algorithms. Notice that we also used the default estimation method in lme(), which is REML. This is different from the estimation method used in the lmer() function (ML). Due to that there will be some differences in the results.

In general, the fixed-effect estimates will be quite similar. The estimates of the variances of the random-effects and the residual variance will be slightly different depending on the estimation method used.

When we are evaluating the error structure, it is advisable to use REML to estimate the model. REML estimation results in less-biased variance estimates than ML.

Autocorrelation

The **nlme** package also includes the `ACF()` function, which allows us to evaluate the correlation structure in the residuals of an `lme()` fitted model.

```
ACF(lme.independence)
```

```
##   lag        ACF
## 1   0  1.0000000
## 2   1 -0.3288128
## 3   2 -0.2932333
## 4   3 -0.2589900
```

This gives the model–based autocorrelations. Autocorrelation is the correlation between elements in a time-series separated by a given interval (called a *lag*). For example, in our model, the autocorrelation for lag 1 (-0.329) indicates that the estimated correlation between residuals at time points that are one wave apart (e.g., Time 1 and Time 2) is -0.329. The estimated correlation between residuals at time points that are two waves apart (e.g., Time 1 and Time 3) is $-.293$. And so on. For more technical information see Pinheiro & Bates (2000).

If the independence assumption is correct, then we expect all autocorrelations to be near zero (NOT to be significantly different than zero). In this case, it seems that most are larger than 0.

We can graphically evaluate whether the autocorrelations are significantly different than 0, by plotting the results of the `ACF()` function.

```
plot(ACF(lme.independence), alpha = 0.05)
```
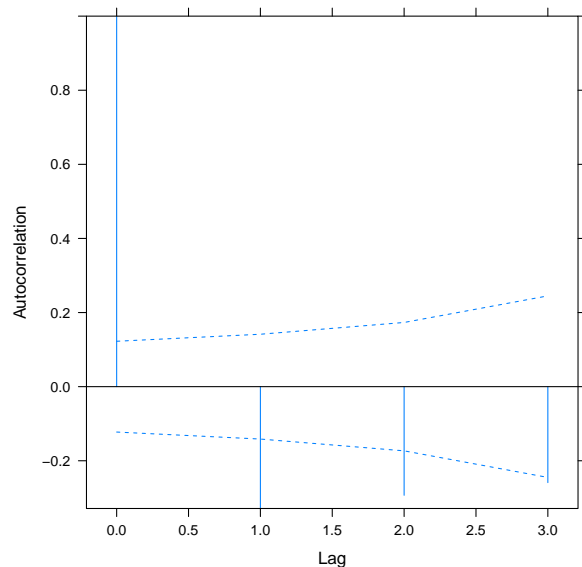

The argument `alpha=` specifies the significance level for approximate two sided critical bounds for the auto-correlation equal to 0. The dashed lines in the plot display the confidence intervals for autocorrelation=0 at each lag. The vertial straight lines are just the magnitude of the model–based autocorrelation at each lag.

The plot suggests that the autocorrelation at lag 1, lag 2, and lag 3 are all significantly different than 0. Hence, we might reject the independence assumption. This plot might also help you to decide what kind of correlational structure you would like to impose on the level-1 errors.

Alternative Level-1 Error Structures

The advantage of the `lme()` function over the `lmer()` function is that it allows us to fit different level-1 error variance–covariance structures. To do this we include the argument `correlation=` in the `lme()` function and specify the structure we want to immpose.

Unstructured

One alternative error structure is unstructured. This error structure allows a different variance for every time point and different covariance between each set of time points. For our four time point example, the within-subjects unstructured variance–covariance matrix is (only the lower half of the matrix is shown since the matrix is symmetric):

$$\mathbf{W_i} = \begin{bmatrix} \sigma_1^2 & & & \\ \sigma_{12} & \sigma_2^2 & & \\ \sigma_{13} & \sigma_{23} & \sigma_3^2 & \\ \sigma_{14} & \sigma_{24} & \sigma_{34} & \sigma_4^2 \end{bmatrix}$$

In this structure we would be estimating four variances and six covariances. We can fit an unstructured error matrix by specifying the argument `correlation=corSymm()`. This tells `lme()` that the variance–covariance structure is symmetric. We also need to use `weights=varIdent(form = ~ 1 | Time)` to allow the variances to differ. `Time` is the variable depicting time in your model. Without this argument, the estimated variance terms on the diagonal will be identical.

```
lme.us = lme(score ~ 1 + grade, random = ~ 1 + grade | id,
                  data = vocab_long, correlation = corSymm(),
                  weights = varIdent(form = ~ 1 | grade) )

summary(lme.us)
```

```
## Linear mixed-effects model fit by REML
##   Data: vocab_long
##        AIC      BIC    logLik
##   894.8793 947.9393 -432.4397
##
## Random effects:
##  Formula: ~1 + grade | id
##  Structure: General positive-definite, Log-Cholesky parametrization
##             StdDev    Corr
## (Intercept) 1.7195693 (Intr)
## grade       0.1511952 -0.048
## Residual    0.7707485
##
## Correlation Structure: General
##  Formula: ~1 | id
##  Parameter estimate(s):
##  Correlation:
##   1      2      3
## 2  0.254
## 3  0.596  0.425
## 4 -0.133  0.034  0.329
## Variance function:
##  Structure: Different standard deviations per stratum
##  Formula: ~1 | grade
```

```
##  Parameter estimates:
##         0        1        2        3
## 1.000000 1.730257 1.720998 1.023303
## Fixed effects: score ~ 1 + grade
##               Value  Std.Error  DF   t-value p-value
## (Intercept) 1.0990125 0.22825865 191  4.814768       0
## grade       0.8063145 0.04853578 191 16.612787       0
##  Correlation:
##       (Intr)
## grade -0.233
##
## Standardized Within-Group Residuals:
##        Min         Q1        Med         Q3        Max
## -2.1546148 -0.3882243  0.1849463  0.7401527  2.4131025
##
## Number of Observations: 256
## Number of Groups: 64
```

Note: Sometimes with the unstructured error matrix, you end up trying to estimate too many parameters given the information (data) provided to the model. In such cases you may get an error letting you knw the estimation did not converge.

To obtain the model–estimated within–subjects error matrix, we use the `getVarCov()` function and include the argument `type = "conditional"`.

```
getVarCov(lme.us, type = "conditional")
```

```
## id 1
## Conditional variance covariance matrix
##          1        2        3        4
## 1  0.594050 0.260740 0.60883 -0.080737
## 2  0.260740 1.778500 0.75247  0.035798
## 3  0.608830 0.752470 1.75950  0.344700
## 4 -0.080737 0.035798 0.34470  0.622060
##    Standard Deviations: 0.77075 1.3336 1.3265 0.78871
```

The model–estimated within–subjects error matrix is

$$\hat{\mathbf{W}}_{\mathbf{i}} = \begin{bmatrix} 0.594 & & & \\ 0.261 & 1.779 & & \\ 0.609 & 0.752 & 1.760 & \\ -0.081 & 0.036 & 0.345 & 0.622 \end{bmatrix}$$

Compound Symmetry

One alternative error structure is compound symmetry. Recall that compound symmetry defines an error structure that has constant variances at each time point and also imposes a non–zero constant covariance among errors at different time points. The within–subjects matrix for compound symmetry is:

$$\mathbf{W}_{\mathbf{i}} = \begin{bmatrix} \sigma_\epsilon^2 & \sigma_c & \sigma_c & \sigma_c \\ \sigma_c & \sigma_\epsilon^2 & \sigma_c & \sigma_c \\ \sigma_c & \sigma_c & \sigma_\epsilon^2 & \sigma_c \\ \sigma_c & \sigma_c & \sigma_c & \sigma_\epsilon^2 \end{bmatrix}$$

In general, the correlation (ρ) between two variables X and Y is just a function of the covariance between them and the standard deviations of those variables,

$$\rho_{XY} = \frac{\mathrm{Cov}(XY)}{\mathrm{SD}(X) \cdot \mathrm{SD}(Y)}$$

Thus the correlation between any two time points when we have compund symmetry is equal to

$$\rho = \frac{\sigma_c}{\sigma_\epsilon^2}$$

since we assume constant variances. Because the covariances are also all equal in the compound symmetry structure, this means that the correlation bewteen every two time points is identical. Sometimes the matrix is wriiten using the correlations rather than the covariances,

$$\mathbf{W_i} = \begin{bmatrix} \sigma_\epsilon^2 & \rho & \rho & \rho \\ \rho & \sigma_\epsilon^2 & \rho & \rho \\ \rho & \rho & \sigma_\epsilon^2 & \rho \\ \rho & \rho & \rho & \sigma_\epsilon^2 \end{bmatrix}$$

We force a compound symmetric error structure by specifying the argument `correlation=corCompSymm()`. We also specify an initial estimate for the ρ parameter. For example, in the syntax below, we fit the model using a compound symmetric structure initiated with $\rho = 0.3$.

```
lme.cs = lme(score ~ 1 + grade, random = ~ 1 + grade | id,
                  data = vocab_long, correlation = corCompSymm(.3))

summary(lme.cs)
```

```
## Linear mixed-effects model fit by REML
##  Data: vocab_long
##        AIC      BIC    logLik
##   891.8282 916.5895 -438.9141
##
## Random effects:
##  Formula: ~1 + grade | id
##  Structure: General positive-definite, Log-Cholesky parametrization
##             StdDev       Corr
## (Intercept) 1.800899414 (Intr)
## grade       0.000141486 0
## Residual    0.945872651
##
## Correlation Structure: Compound symmetry
##  Formula: ~1 | id
##  Parameter estimate(s):
##         Rho
## -0.007237347
## Fixed effects: score ~ 1 + grade
##                 Value  Std.Error  DF  t-value p-value
## (Intercept) 1.4132969 0.24582669 191  5.74916       0
## grade       0.7466562 0.05306689 191 14.07010       0
##  Correlation:
##       (Intr)
## grade -0.324
```

```
##
## Standardized Within-Group Residuals:
##         Min          Q1         Med          Q3         Max
## -2.33858087 -0.66030554 -0.03241567  0.64468155  2.65744599
##
## Number of Observations: 256
## Number of Groups: 64
```

In the output, we see that the final model–based estimate of ρ is $\hat{\rho} = -.007$. From the output we can also obtain the estimated error variance (0.895). Computing the covariance estimate, we get $\sigma_c = \rho \sigma_\epsilon^2 = -.007(0.895) = -.006$. The model-estimated within-subjects error matrix is then,

$$\hat{\mathbf{W}}_{\mathbf{i}} = \begin{bmatrix} 0.895 & & & \\ -0.006 & 0.895 & & \\ -0.006 & -0.006 & 0.895 & \\ -0.006 & -0.006 & -0.006 & 0.895 \end{bmatrix}$$

This can be verfied using R.

```
getVarCov(lme.cs, type = "conditional")
```

```
## id 1
## Conditional variance covariance matrix
##            1          2          3          4
## 1  0.8946800 -0.0064751 -0.0064751 -0.0064751
## 2 -0.0064751  0.8946800 -0.0064751 -0.0064751
## 3 -0.0064751 -0.0064751  0.8946800 -0.0064751
## 4 -0.0064751 -0.0064751 -0.0064751  0.8946800
##   Standard Deviations: 0.94587 0.94587 0.94587 0.94587
```

Compound Symmetry ($\rho = 0.5$)

If you specify a different starting value for ρ other than 0.3, it is likely that you will get a different final estimate.

```
lme.cs2 = lme(score ~ 1 + grade, random = ~ 1 + grade | id,
                  data = vocab_long, correlation = corCompSymm(.5))
```

```
summary(lme.cs2)
```

```
## Linear mixed-effects model fit by REML
##   Data: vocab_long
##        AIC      BIC    logLik
##   891.8282 916.5895 -438.9141
##
## Random effects:
##   Formula: ~1 + grade | id
##   Structure: General positive-definite, Log-Cholesky parametrization
##             StdDev        Corr
## (Intercept) 1.7663863306 (Intr)
## grade       0.0001347917 0
## Residual    1.0089111648
##
## Correlation Structure: Compound symmetry
##   Formula: ~1 | id
##   Parameter estimate(s):
```

```
##        Rho
## 0.1147081
## Fixed effects: score ~ 1 + grade
##                 Value  Std.Error  DF   t-value p-value
## (Intercept) 1.4132969 0.24583024 191  5.749077       0
## grade       0.7466563 0.05306659 191 14.070175       0
##  Correlation:
##       (Intr)
## grade -0.324
##
## Standardized Within-Group Residuals:
##          Min          Q1         Med          Q3         Max
## -2.33033872 -0.62209624 -0.03938941  0.58199975  2.49993154
##
## Number of Observations: 256
## Number of Groups: 64
#getVarCov(lme.cs2, type = "conditional")
```

As you see in the output, the final estimated value is now $\hat{\rho} = .114$. The value you specify is a personal judgement. The model–estimated within–subjects error matrix from this model is:

$$\hat{\mathbf{W}}_i = \begin{bmatrix} 1.018 & & & \\ 0.117 & 1.018 & & \\ 0.117 & 0.117 & 1.018 & \\ 0.117 & 0.117 & 0.117 & 1.018 \end{bmatrix}$$

Compound Symmetry ($\rho = 0$)

Finally, if you do not specify an initial estimate fro ρ, the starting values is taken as 0 by default.

```
lme.cs3 = lme(score ~ 1 + grade, random = ~ 1 + grade | id,
                data = vocab_long, correlation = corCompSymm())

summary(lme.cs3)

## Linear mixed-effects model fit by REML
##  Data: vocab_long
##        AIC      BIC    logLik
##   891.8282 916.5895 -438.9141
##
## Random effects:
##  Formula: ~1 + grade | id
##  Structure: General positive-definite, Log-Cholesky parametrization
##             StdDev        Corr
## (Intercept) 1.8346433781 (Intr)
## grade       0.0001168135 0
## Residual    0.8786944919
##
## Correlation Structure: Compound symmetry
##  Formula: ~1 | id
##  Parameter estimate(s):
##        Rho
## -0.1671231
## Fixed effects: score ~ 1 + grade
```

```
##              Value Std.Error  DF   t-value p-value
## (Intercept) 1.4132969 0.2458302 191  5.749078       0
## grade       0.7466562 0.0530666 191 14.070173       0
##  Correlation:
##      (Intr)
## grade -0.324
##
## Standardized Within-Group Residuals:
##         Min          Q1         Med          Q3         Max
## -2.35990447 -0.69196529 -0.02273795  0.64032599  2.85087473
##
## Number of Observations: 256
## Number of Groups: 64
```
```
#getVarCov(lme.cs3, type = "conditional")
```

The model-estimated within–subjects error matrix from this model is:

$$
\hat{\mathbf{W}}_{\mathbf{i}} = \begin{bmatrix}
0.772 & & & \\
-0.129 & 0.772 & & \\
-0.129 & -0.129 & 0.772 & \\
-0.129 & -0.129 & -0.129 & 0.772
\end{bmatrix}
$$

Compound Symmetry ($\rho = 0.5$, Fixed)

If you have theoretical reason to do so, you can also fix the correlation at the given value by specifying the ρ value and also including the argument fixed=TRUE. This forces the final estimate of ρ to be whatever value is specified.

```
lme.cs4 = lme(score ~ 1 + grade, random = ~ 1 + grade | id,
              data = vocab_long, correlation = corCompSymm(.5, fixed = TRUE))

summary(lme.cs4)
```

```
## Linear mixed-effects model fit by REML
##  Data: vocab_long
##       AIC      BIC    logLik
##   889.8282 911.0522 -438.9141
##
## Random effects:
##  Formula: ~1 + grade | id
##  Structure: General positive-definite, Log-Cholesky parametrization
##             StdDev        Corr
## (Intercept) 1.52826505400 (Intr)
## grade       0.00006548757 0
## Residual    1.34249892512
##
## Correlation Structure: Compound symmetry
##  Formula: ~1 | id
##  Parameter estimate(s):
## Rho
## 0.5
## Fixed effects: score ~ 1 + grade
##                 Value  Std.Error  DF   t-value p-value
## (Intercept) 1.4132969 0.24582616 191  5.749172       0
```

```
## grade        0.7466563 0.05306693 191 14.070086        0
##  Correlation:
##       (Intr)
## grade -0.324
##
## Standardized Within-Group Residuals:
##        Min         Q1         Med         Q3         Max
## -2.41089146 -0.48124469 -0.04449588  0.47653617  2.36389468
##
## Number of Observations: 256
## Number of Groups: 64
```
```
#getVarCov(lme.cs4, type = "conditional")
```

Here, the model–estimated within–subjects error matrix from this model is:

$$
\hat{\mathbf{W_i}} = \begin{bmatrix}
1.802 & & & \\
0.901 & 1.802 & & \\
0.901 & 0.901 & 1.802 & \\
0.901 & 0.901 & 0.901 & 1.802
\end{bmatrix}
$$

Comparison of the Compound Symmetry Models

	Compund Symmetric Models			
	$\rho = 0.3$	$\rho = 0.5$	$\rho = 0$	$\rho = 0.5$, Fixed
Fixed-Effects				
Grade	0.747***	0.747***	0.747***	0.747***
	(0.053)	(0.053)	(0.053)	(0.053)
Constant	1.413***	1.413***	1.413***	1.413***
	(0.246)	(0.246)	(0.246)	(0.246)
Variance Components				
$\hat{\sigma}^2_{b_{0i}}$	3.243	3.120	3.366	2.336
$\hat{\sigma}^2_{b_{1i}}$	0.000000020	0.000000018	0.000000014	0.000000004
$\hat{\sigma}^2_{\epsilon}$	0.895	1.018	0.772	1.802

Note. *p<0.1; **p<0.05; ***p<0.01

In all the models fitted, the fixed–effects are essentially the same, regardless of the ρ value. The estimated error variance, and the variance of the random–effects depends on the covariances (or correlation) that is estimated.

Autoregressive Structures

Autoregressive structures specify that the correlation at the same lag values are correlated equally. In the *first order autoregressive* structure, AR(1), the correlation between any two residuals e_i and e_j is a power of a single correlation parameter (ρ) and the ordination of measurement occasions (the lag). That is,

$$\text{Corr}(e_i, e_j) = \rho^{|i-j|}$$

This implies that the larger the lag, the weaker the serial correlation. The AR(1) error structure is

$$\mathbf{W_i} = \begin{bmatrix} \sigma_\epsilon^2 & & & \\ \rho\sigma_\epsilon^2 & \sigma_\epsilon^2 & & \\ \rho^2\sigma_\epsilon^2 & \rho\sigma_\epsilon^2 & \sigma_\epsilon^2 & \\ \rho^3\sigma_\epsilon^2 & \rho^2\sigma_\epsilon^2 & \rho\sigma_\epsilon^2 & \sigma_\epsilon^2 \end{bmatrix}$$

Not only do time points separated by the same lag have the same covariance (correlation), but the correlation in each diagonal band is less than (or equal to) the diagonal bands above it and greater than (or equal to) the diagonal bands below it.

$$\rho \leq \rho^2 \leq \rho^3$$

To fit this structure, we use `correlation=corAR1()`.

```
lme.ar1 = lme(score ~ 1 + grade, random = ~ 1 + grade | id,
                data = vocab_long, correlation = corAR1())

summary(lme.ar1)
```

```
## Linear mixed-effects model fit by REML
##  Data: vocab_long
##        AIC      BIC    logLik
##   891.7228 916.4841 -438.8614
##
## Random effects:
##  Formula: ~1 + grade | id
##  Structure: General positive-definite, Log-Cholesky parametrization
##             StdDev       Corr
## (Intercept) 1.8053721088 (Intr)
## grade       0.0001153592 0
## Residual    0.9407654576
##
## Correlation Structure: AR(1)
##  Formula: ~1 | id
##  Parameter estimate(s):
##         Phi
## -0.03545799
## Fixed effects: score ~ 1 + grade
##                  Value  Std.Error  DF   t-value p-value
## (Intercept) 1.4200026 0.24558142 191  5.782207       0
## grade       0.7448768 0.05209431 191 14.298621       0
##  Correlation:
##       (Intr)
## grade -0.318
```

```
##
## Standardized Within-Group Residuals:
##          Min          Q1         Med          Q3         Max
## -2.34403652 -0.67975431 -0.03245168  0.64417801  2.65456260
##
## Number of Observations: 256
## Number of Groups: 64
```

The estimated ρ for the first order autoregressive error structure is -0.035. We also see that the residual variance estimate is $0.941^2 = 0.885$. Thus,

$$
\hat{\mathbf{W}}_\mathbf{i} = \begin{bmatrix} 0.885 & & & \\ (-0.035)(0.885) & 0.885 & & \\ (-0.035)^2(0.885) & (-0.035)(0.885) & 0.885 & \\ (-0.035)^3(0.885) & (-0.035)^2(0.885) & (-0.035)(0.885) & 0.885 \end{bmatrix}
$$
$$
= \begin{bmatrix} 0.885 & & & \\ -0.031 & 0.885 & & \\ -0.0011 & -0.031 & 0.885 & \\ -0.00004 & -0.0011 & -0.031 & 0.885 \end{bmatrix}
$$

We can again get the same result directly using R

```
# Obtain estimated W_i matrix
getVarCov(lme.ar1, type = "conditional")
```

```
## id 1
## Conditional variance covariance matrix
##              1          2          3            4
## 1  0.885040000 -0.0313820  0.0011127 -0.000039455
## 2 -0.031382000  0.8850400 -0.0313820  0.001112700
## 3  0.001112700 -0.0313820  0.8850400 -0.031382000
## 4 -0.000039455  0.0011127 -0.0313820  0.885040000
##    Standard Deviations: 0.94077 0.94077 0.94077 0.94077
```

As a practical note, the AR(1) structure only makes sense if the measurement occasions are equally spaced.

Other Structures

There are several other error structures as well (e.g., Toeplitz, banded diagonal, banded two). Most of the textbooks recommended for the course discuss several of these structures, including Fitzmaurice, Laird, & Ware (2011), Pinheiro & Bates (2000), and Singer & Willett (2003).

Comparing Error Structures

To compare models fitted with different level–1 error structures, we can compare the information criteria. In our example,

```r
aictab(
  cand.set = list(lme.independence, lme.us, lme.cs, lme.ar1),
  modnames = c("Independent", "Unstructured", "Compound Symmetry", "AR(1)")
  )
```

```
## Warning in aictab.AIClme(cand.set = list(lme.independence, lme.us, lme.cs, :
## Model selection for fixed effects is only appropriate with method=ML:
## REML (default) should only be used to select random effects for a constant set of fixed effects

##
## Model selection based on AICc:
##
##                      K  AICc Delta_AICc AICcWt Cum.Wt  Res.LL
## Independent          6 890.17       0.00   0.57   0.57 -438.91
## AR(1)                7 892.17       2.01   0.21   0.78 -438.86
## Compound Symmetry    7 892.28       2.11   0.20   0.98 -438.91
## Unstructured        15 896.88       6.71   0.02   1.00 -432.44
```

The empirical evidence supports three level–1 residual structures: independent, first–order autoregressive, and compound symmetry. I would probably adopt the independent structure as it is the most parsimonious of the error structures.

References

Crowder, M., & Hand, D. (1990). *Analysis of repeated measures*. London: Chapman; Hall.

Fitzmaurice, G. M., Laird, N. M., & Ware, J. H. (2011). *Applied longitudinal analysis* (2nd ed.). New York: Wiley.

Kohli, N. (2016). Random coefficient models: Error structures. Class notes.

Long, J. D. (2012). *Longitudinal data analysis for the behavioral sciences using R*. Thousand Oaks, CA: Sage Publications, Inc.

Pinheiro, J. C., & Bates, D. (2000). *Mixed-effects models in S and S-PLUS*. New York: Springer.

Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. New York: Oxford University Press.