# Intraclass Correlation Coefficient in mixed model with random slopes

Asked 8 years, 8 months ago Modified 5 years, 5 months ago Viewed 12k times

16

I have the following model `m_plot` fitted with `lme4::lmer` with crossed random effects for participants ( `lfdn` ) and items ( `content` ):

```
Random effects:
 Groups   Name             Variance Std.Dev. Corr
 lfdn     (Intercept)      172.173  13.121
          role1             62.351   7.896    0.03
          inference1        24.640   4.964    0.08 -0.30
          inference2        52.366   7.236   -0.05  0.17 -0.83
          inference3        21.295   4.615   -0.03  0.22  0.86 -0.77
 content  (Intercept)       23.872   4.886
          role1              2.497   1.580   -1.00
          inference1        18.929   4.351    0.52 -0.52
          inference2        14.716   3.836   -0.16  0.16 -0.08
          inference3        17.782   4.217   -0.17  0.17  0.25 -0.79
          role1:inference1   9.041   3.007    0.10 -0.10 -0.10 -0.21  0.16
          role1:inference2   5.968   2.443   -0.60  0.60 -0.11  0.78 -0.48 -0.50
          role1:inference3   4.420   2.102    0.30 -0.30  0.05 -0.97  0.71  0.37 -0.90
 Residual                   553.987  23.537
Number of obs: 3480, groups:  lfdn, 435 content, 20
```

I want to know the Intraclass Correlation Coefficients (ICC) for participants and items. Thanks to this great answer I in principle know how to get the ICC for my model. However, I am unsure on whether or not to include the random slopes or not:

```
vars <- lapply(summary(m_plot)$varcor, diag)
resid_var <- attr(summary(m_plot)$varcor, "sc")^2
total_var <- sum(sapply(vars, sum), resid_var)

# with random slopes
sapply(vars, sum)/total_var
##       lfdn    content
## 0.33822396 0.09880349

# only random intercepts:
sapply(vars, function(x) x[1]) / total_var
##   lfdn.(Intercept) content.(Intercept)
##         0.17496587          0.02425948
```

What is the appropriate measure for the correlation between two responses from the same participant respective to the same item?

mixed-model lme4-nlme intraclass-correlation

Share Cite Improve this question Follow

edited Apr 13, 2017 at 12:44
Community Bot
1

asked Sep 15, 2014 at 17:11
Henrik
13.9k 10 67 128

1 Merlo et al 2005 "A brief conceptual tutorial on multilevel analysis in social epidemiology: investigating contextual phenomena in different groups of people" might be a useful reference. – N Brouwer Nov 13, 2014 at 15:26

@Henrik did you ever find an answer to this question? I'm interested as well. – Patrick S. Forscher Dec 8, 2017 at 19:22

2 @PatrickS.Forscher As far as I understand, ICC does not make sense with random slopes. I have learned this from Jake Westfall. – Henrik Dec 10,

2017 at 22:20

Got a link to a reading relevant reading by chance? – Patrick S. Forscher Dec 11, 2017 at 3:23

1 @PatrickS.Forscher As you can see, Jake Westfall now provided a great answer. – Henrik Dec 12, 2017 at 15:31

## 1 Answer

Sorted by: Highest score (default)

14

Basically there's no single number or estimate that can summarize the degree of clustering in a random slopes model.

The intra-class correlation (ICC) can only be written as a simple proportion of variances in random-intercepts-only models. To see why, a sketch of the derivation of the ICC expression can be found here.

When you throw random slopes into the model equation, following the same steps leads instead to the ICC expression on page 5 of this paper. As you can see, that complicated expression is a function of the predictor X. To see more intuitively why var(Y) depends on X when there are random slopes, check out page 30 of these slides ("Why does the variance depend on x?").

Because the ICC is a function of the predictors (the x-values), it can only be computed for particular sets of x-values. You could perhaps try something like reporting the ICC at the joint average of the x-values, but this estimate will be demonstrably inaccurate for the majority of the observations.

Everything I've said still only refers to cases where there is a single random factor. With multiple random factors it becomes even more complicated. For example, in a multi-site project where participants at each site respond to a sample of stimuli (i.e., 3 random factors: site, participant, stimulus), we could ask about many different ICCs: What is the expected correlation between two responses at the same site, to the same stimulus, from different participants? How about at different sites, the same stimulus, and different participants? And so on. @rvl mentions these complications in the answer that the OP linked to.

So as you can see, the only case where we can summarize the degree of clustering with a single value is the single-random-factor random-intercept-only case. Because this is such a small proportion of real-world cases, ICCs are not that useful most of the time. So my general recommendation is to not even worry about them. Instead I recommend just reporting the variance components (preferably in standard deviation form).

Share Cite Improve this answer Follow

edited Dec 12, 2017 at 17:34

answered Dec 12, 2017 at 14:25

Jake Westfall
**12.1k** 2 51 100