**Assignment 4: Linear Mixed Effects Models**

Steven Jackson

Dalhousie University

Department of Psychiatry

Dr. Igor Yakovenko

April 26th, 2024

This dataset contains repeated-measures self-report of extraversion, neuroticism, and satisfaction of life among 263 participants over a maximum of 20 days. The variables tipm.E and tipm.N refer to extraversion and neuroticism, respectively, while swl refers to satisfaction with life, day refers to any day of a possible 20 ranging from 1-21 (day 0, baseline, used a different questionnaire and is not included), and id is a unique identifier assigned to each participant. The results below detail how the dataset was prepared and analyzed, complete with APA tables, descriptive statistics, and model comparison.

We started by cloning the repository from http://github.com/iyakoven/PSYR6003~Assignment~4 . Next, the dataset was imported, and missing values removed. Because the dataset included many unneeded variables, only the id, day, swl, tipm.E, and tipm.N variables were selected. The raw data was then plotted, without accounting for repeated measurements, to look at the relationship between extraversion and satisfaction with life, as well as neuroticism and satisfaction with life. The relationship appeared linear. Upon visualizing the univariate distributions, extraversion appears normally distributed. Satisfaction with life and neuroticism appear a little less so, the former looking a little left-skewed, and the latter right-skewed. However, the skew is not severe, and in the case of extraversion and neuroticism there are only six bins in the histogram, which may not provide the precision necessary to see a neatly normal distribution. We will consider these assumptions met in the absence of any other confounding elements.

The mean extraversion score in the sample was 4.18 (SD = 1.52), and the mean neuroticism score was 3.49 (SD = 1.54). The two items are correlated at -0.33 [-0.36, -0.31]. These descriptive statistics are shown in Table 1 below.

**Table 1**

*Means, standard deviations, and correlations with confidence intervals*

|  |  |  |  |
| --- | --- | --- | --- |
| Variable | *M* | *SD* | 1 |
|  |  |  |  |
| 1. tipm.E | 4.18 | 1.52 |  |
|  |  |  |  |
| 2. tipm.N | 3.49 | 1.54 | -.33\*\* |
|  |  |  | [-.36, -.31] |
|  |  |  |  |

*Note.* *M* and *SD* are used to represent mean and standard deviation, respectively. Values in square brackets indicate the 95% confidence interval for each correlation. The confidence interval is a plausible range of population correlations that could have caused the sample correlation (Cumming, 2014). \* indicates *p* < .05. \*\* indicates *p* < .01.

Next, a simple mixed linear model was prepared with no predictors and the intercept set to random. Our ICC is ~0.74, indicating that a high degree of the variance of life satisfaction is due to clustering (i.e., the participant ID). This indicates that we are correct in using a mixed effects model, rather than a regular linear regression model. In fact, if we were to treat all observations as independent (i.e., not use a mixed effects model), we would artificially increase our sample size more than 12-fold. It is crucial to use a linear mixed model with these data.

Hypothesis #1 and #2 state that extraversion will be positively associated with satisfaction with life, and neuroticism will be negatively associated with quality of life, respectively. For hypotheses #1 and #2, we will create a model for satisfaction with life using extraversion as fixed (i.e., extraversion predicts satisfaction with life and this effect is the same across participants), and one with extraversion as a random effect (i.e., extraversion predicts satisfaction with life to a varying degree across participants). We will select the parameter with the best fit, then repeat this process for neuroticism in order to test hypothesis #2. We used ML estimators to allow for the comparison of nested models.

The comparison between the two models tests the two equations below against each other:

*Random effects model*

*Fixed effects model*

After comparing the fixed and random extraversion models, we can see that using extraversion as a random effect is the best fit. The random effect model had a very high Bayes factors (decisive evidence for using random instead of fixed), and smaller AIC & BIC than the fixed model, both indicating better fits. According to the available metrics, extraversion should be used as a random effect.

Next, we will take the random-effect extraversion model and add neuroticism. Since extraversion is already known to be better as a random effect, we'll set extraversion to random, try neuroticism as random & fixed in this combined model, and again pick the best fitting parameter based on the fit metrics.

After doing the comparisons, the best-fitting models are random extraversion and random neuroticism. The model comparison above shows that the random-random model has a lower AIC & BIC, as well as a Bayes factor denoting decisive evidence in favour of the random-random model. Lastly, the R^2 changes for the intercept and residual is -0.28 and -0.08 respectively, indicating a 28% increase in predicted variance around the fixed mean, and an 8% increase in the predicted variance of the cluster mean using the random model instead of the fixed one. All of these metrics point to the random effects model having a better fit than the mixed model.

For the third and final hypothesis, we will test whether the effects are similar both within-participants and between-participants. We can do this by visualizing our existing model. Upon visualization, the diagnostics for the random effects model look very strong. The residuals are clearly normally distributed, the residual dependence plot is flat, and the S-L plot is flat as well.

As seen in the generated graphs, the relationship between satisfaction with life, extraversion, and neuroticism is similar between and within participants. This confirms hypothesis #3. The intercept R^2 is negative, indicating that the model predicted none of the variance around the overall mean satisfaction with life. The residual R^2, however, was 0.26, indicating that the model predicted ~26% of the variance around the mean for each participant. The random effect model we chose is summarized in Table 2, while the fixed effect model is shown in Table 3.

**Table 2**

*Random effect linear model results using satisfaction with life as the criterion*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Predictor | *Variance* | *Standard Deviation* | Residual R2 | | ICC | |
| (Intercept) | 2.02 | 1.42 |  | |  | |
| tipm.E | 0.020 | 0.14 |  | |  | |
| tipm.N | 0.035 | 0.19 |  | |  | |
|  |  |  | 0.261 | | 0.74 | |
|  |  |  | |  | |  | |

*Note*. *b* represents unstandardized regression weights.

**Table 3**

*Fixed effect regression results using satisfaction with life as the criterion*

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Predictor | | *b* | | | *SE* | | Fit | | | |
| (Intercept) | | 4.50 | | | 0.12 | |  | | | |
| tipm.E | | 0.27 | | | 0.02 | |  | | | |
| tipm.N | | -0.38 | | | 0.02 | |  | | | |
|  | |  | | |  | | *Conditional R2: 0.789*  *Marginal R2: 0.094* | | | |
|  |  | |  |  | |  | |  | |  | |
|  | |  | | |  | | | |

*Note. b* represents unstandardized regression weights.

In conclusion, extraversion positively predicts satisfaction with life and neuroticism negatively predicts satisfaction with life, confirming hypotheses #1 and #2. The effect is very similar within and between clusters. In lay terms, this means both that if a person becomes more extraverted, they also become more satisfied with life, and that people who are more extraverted than other people are also more likely to be satisfied with life than those people; this confirms hypothesis #3.