**Assignment 4: Linear Mixed Effects Modelling**

Mikaela Ethier-Gagnon

Department of Psychiatry, Dalhousie University

PSYR6003: Fundamentals of Applied Statistics

Dr. Igor Yakovenko

April 23rd, 2024

**Assignment 4: Linear Mixed Effects Modelling**

In the present study, *N*=236 participants were measured with a daily questionnaire for up to 20 days for a total of 4252 observations. Participants were assessed for satisfaction with life (SWL), extraversion, and neuroticism. Items ranged from 1 ‘strongly disagree’ to 7 ‘strongly agree’ for all scales, and component items for each measure were averaged into a total score such that higher number indicate higher SWL, extraversion, or neuroticism, respectively. This dataset was obtained from Dr. Igor Yakovenko’s repository on GitHub <https://github.com/iyakoven/PSYR6003-Assignment-4>. We hypothesized that extraversion would be positively associated with SWL, whereas neuroticism will be negatively associated with SWL. Moreover, we anticipated effects to be similar within participants over time, and between participants. Data analysis was completed using a custom R script in R Studio (version 4.3.2; R Core Team, 2023).

Descriptive statistics were calculated using the *apaTables* package (Stanley, 2021). Outcomes of interest included SWL (*M=*4.43, *SD*=1.61), extraversion (*M=*4.18, *SD*=1.52), and neuroticism (*M=*3.49, *SD*=1.54). Means, standard deviations, and bivariate correlations are presented in Table 1. As hypothesized, SWL was found to be weakly positively correlated with extraversion (*r*=0.38, *p* < .01), whereas neuroticism was moderately negatively correlated (*r*=-0.45, *p* < .01). Neuroticism and extraversion were also observed to be weakly negatively correlated with one another (*r*=-0.33, *p* < .01).

Given that the data was clustered for each participant over time, hypotheses were tested using linear mixed effect models, with SWL specified as the outcome variable, and extraversion and neuroticism as predictors. Univariate distributions of SWL, extraversion, and neuroticism were observed to be relatively normal, with a slight negative skew for neuroticism, and a minor positive skew for SWL. Given that the data was normally distributed, to determine whether predictor variables would be specified as random, fixed, or both, we conducted nested model comparisons using the *lme4* package(Bates et al., 2015). To determine the use of fixed vs. random slopes, we conducted nested model comparisons using maximum likelihood estimation. The model which best fit the data included participant ID as the cluster variable and neuroticism, and extraversion specified as both fixed and random variables. Thus, for the final model, we proceeded the analysis with both fixed and random slopes, using maximum likelihood estimation with the following algebraic notation: SWL = β00 + β10(Extraversion) + β20(Neuroticism) + r1(extraversion) +r2(neuroticism) + r0 + e. Statistical assumptions of the model including normality, linearity, and homoscedasticity were tested using the *Flexplot* package (Fife, 2024). Residuals were revealed to be normally distributed; linearity was also observed to be met as evidenced by a flat line of best fit in the Residual Dependence Plot. However, the assumption of homoscedasticity was violated as the line of best fit presented in the SL plot was not flat. To address this violated assumption, there are several different approaches including transformation of the dependent variable, weighted least squares, robust standard errors and generalized least squares. However, for the purposes of this analysis this assumption violation was ignored. Results of the model, including estimates, confidence intervals, and standard errors for fixed and random effects, as well as *R2* values, are found in Table 2. The model revealed an intraclass correlation coefficient of 0.743, indicating that 74.3% of the variance in SWL can be accounted for by clustering, suggesting that linear mixed effects modeling is an appropriate approach for this analysis. Consistent with our hypothesis, extraversion was found to be positively associated with SWL, specifically, with every one-point increase in extraversion, there is a 0.16-point increase in SWL. Further, as anticipated, neuroticism was negatively associated with SWL, notably, a 1-point increase in neuroticism was associated with a -0.21-point decrease in SWL. However, contrary to our hypothesis, effects for level 1 (within participants over time) and level 2 (between participants) differed. Specifically, we did not predict any variability of SWL around the overall mean of SWL for all participants (i.e., between subjects), however, we predicted 26% of the variance in SWL from day to day for each participant (i.e., within-subjects). When further investigating random effects, the variance attributable to the individual participants was 2.02, when dividing this by the total variance, we achieve 95%. This means that merely 95% of the variance can be explained by differences between individuals, which is leftover after the variance explained by fixed effects. Fixed effects alone accounted for 9% of the variance in SWL, whereas the whole model including both fixed and random effects accounted for 79% of the variance in SWL.

To conclude, consistent with our hypothesis, extraversion was positively associated with SWL, and neuroticism was negatively associated with SWL. However, contrary to our hypothesis, effects at level 1 and level 2 differed, as 26% of the variance in SWL can be explained at level 1 whereas between-subjects level 2 effects did not explain any variance in SWL.

Table 1

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

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Variable | *M* | *SD* | 1 | 2 |
|  |  |  |  |  |
| 1. Extraversion | 4.18 | 1.52 |  |  |
|  |  |  |  |  |
| 2. Neuroticism | 3.49 | 1.54 | -.33\*\* |  |
|  |  |  | [-.36, -.31] |  |
|  |  |  |  |  |
| 3. Satisfaction with Life | 4.43 | 1.61 | .38\*\* | -.45\*\* |
|  |  |  | [.35, .41] | [-.47, -.43] |
|  |  |  |  |  |

*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.

Table 2

*Mixed effect model results using satisfaction with life as the criterion*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Predictor | | Fixed Effects | | | |
|  | *B* | | *B*  95% CI  [LL, UL] | *SE* | Fit |
| (Intercept) | 4.51\*\* | | [4.27, 4.74] | 0.12 |  |
| Extraversion | 0.16\*\* | | [0.13, 0.19] | 0.02 |  |
| Neuroticism | -0.21\*\* | | [-0.24, -0.18] | 0.02 |  |
|  |  | |  |  | *R2* Intercept / *R2* Residual = -0.06 / 026 |
|  |  | |  |  | *ICC* = 0.74 |
| Predictor |  | | Random Effects | | |
|  | *Variance* | | *SD* |  | Fit |
| (Intercept) | 2.03 | | 1.42 |  |  |
| Extraversion | 0.02 | | 0.14 |  |  |
| Neuroticism | 0.04 | | 0.19 |  |  |
| Residual | 0.49 | | 0.70 |  |  |
|  |  | |  |  | Marginal *R2 /* Conditional *R2* = 0.09 / 0.89 |
|  |  | |  |  |  |  |

*Note. B* represents the fixed effect estimates. *SE* represents the standard errors of the estimates. *LL* and *UL* indicate the lower and upper limits of a confidence interval, respectively. *R2* represents the coefficient of determination. *ICC* represents the intraclass correlation coefficient. *SD* represents the standard deviation.   
\* indicates *p* < .05. \*\* indicates *p* < .01.

**References**

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1-48. <https://doi.org/10.18637/jss.v067.i01>

Fife, D. (2024). *flexplot: Graphically based data analysis using 'flexplot'*. R package version 0.20.3.

R Core Team. (2023). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

Stanley, D. (2021). *apaTables: Create American Psychological Association (APA) style tables*. R package version 2.0.8. https://CRAN.R-project.org/package=apaTables>.