**PSYR 6003 Assignment 2: The General Linear Model**

Holly Carbyn

Department of Psychiatry, Dalhousie University

PSYR 6003: Fundamentals of Applied Statistics

Dr. Igor Yakovenko

April 7th, 2024

**Assignment 1: The General Linear Model**

In this data analysis, we evaluated a dataset (*n* = 137) using statistical software R version 4.3.2 (Eye Holes). This dataset consisted of participant scores on personability traits using the Multidimensional Perfectionism Scale Short Form: Socially Prescribed Subscale (3 year version), the PANAS: Negative Affect (3 year version), and the Ten Item Personality Measure: Conscientiousness Subscale (3-year version). The dataset also includes sex as a variable (asks the sex of the participant). We used the ‘Tidyverse’ package (Wickham et al., 2019) to conduct analyses. The dataset was cleaned by removing people who chose “other” for sex, and by removing cases with missing values (*n* = 131). We calculated subscale totals for socially prescribed perfectionism (SPP), conscientiousness, and negative affect for each participant by taking the mean of all the constituent items in each of the three scales.

**Descriptive Statistics & Bivariate Correlations**

Descriptive statistics and bivariate correlations were calculated by creating a correlation table using the ‘apaTables’ package (Stanley, 202) for our variables of interest; SPP (*M* = 4.38, *SD* = 1.43), conscientiousness (*M* = 5.03, *SD* = 1.29), negative affect (*M* = 2.44, *SD* = 0.98), and sex (*M* = 0.15, *SD* = 0.35), as seen in *Table 1*. Bivariate correlations were calculated to examine the relationships among these variables. SPP was positively correlated with negative affect (*r* = 0.37, *CI* = [0.21,0.51], *p* < 0.01) and negatively correlated with conscientiousness (*r* = -0.21, *CI* = [-0.37, -0.04], *p* < 0.05). Additionally, there was a significant negative correlation between conscientiousness and negative affect (*r* = -0.37, CI = [-.51, -.22], p < 0.01). No significant correlation was found between sex and any other variable of interest (see *Table 1*). These findings suggest that SPP is associated with higher levels of negative affect and lower levels of conscientiousness, while conscientiousness is inversely related to negative affect.

**Multiple Regression**

We ran a multiple regression analysis to see whether 1) sex, conscientiousness, and SPP all significantly predict negative affect and 2) SPP predicts unique variance in negative affect over and above sex and conscientiousness in a meaningful way.

We first visualized the univariate distributions of our analysis variables by creating graphs and plots using the ‘flexplot’ package (Phillipp, 2021). We see that scores for SPP have a slightly negatively skewed distribution, conscientiousness has a negatively skewed distribution, and negative affect has a slightly positively skewed distribution. There are many more females than males in the dataset, which is evident from the graph generated for the ‘sex’ variable. The model visualization for hypothesis one (H1) suggests that conscientiousness, SPP, and sex may possibly influence negative affect. The added variable plot (AVP) visualization for the full model of hypothesis two (H2) suggests that SPP may potentially predict negative affect above and beyond conscientiousness and sex.

We next used the package ‘flexplot’ (Phillipp, 2021) to create graphs and plots to check the statistical assumptions for the multiple regression models. A Histogram of our residuals showed a slightly negatively skewed distribution. This led us to check for normality of residuals by running a sensitivity analysis with robust SEs using the MASS package (Venables & Ripley, 2002). The robust model showed similar slopes to the original model, so we ignored the violation of normality of residuals. The assumption of linearity may be violated. The Residual Dependence plot shows that the tail ends of the data curve upwards, however, this could be due to high leverage data points/outliers. To combat this violation of linearity, one could add a polynomial variable (which is outside the scope of this class). The assumption of homogeneity of variance or ‘homoscedasticity’ is violated due to the small positive slope of the S-L plot. We ignored this slight violation, but one could fix it by applying transformations like logarithm, square root, or inverse to the dependent variable to stabilize variance across different levels of the independent variables. The assumption of independence of observations is met by design as we are not evaluating repeated measures or participants who are grouped or related. Under these circumstances we can assume independence.

For H1, we see that the SPP, conscientiousness, and sex predict about 27% of variance in negative affect (*r2* = 0.269, 95% CI[0.13, 0.37]). Looking at the semi-partial R2, we can see that conscientiousness explains about 16%, SPP explains about 8.4% and sex explains about 2.5% of the variance. This means that conscientiousness explains most of the variance in negative affect. All three of the estimates for SPP, conscientiousness, and sex have a 95% CI that does not cross 0, which means they are having an effect. For every one point increase in conscientiousness and SPP scores, negative affect scores go down by 0.26 points and up by 0.2 points respectively (*Table 2*). The predicted differences between the full model (with the three predictors) versus the reduced model (no predictors) suggest that the full model can potentially predict up to a 0.64 point differential in negative affect, which is a fair amount given its range, further cementing the difference between the models.

For H2, the fit indices suggest that the full model (Negative Affect = b0 + b1\*sex + b2\*conscientiousness + b3\*SPP + e) is a better fit than the reduced model (Negative Affect = b0 + b1\*sex + b2\*conscientiousness + e). This is evident because the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) values are both much smaller for the full model (*AIC* = 333.26, *BIC* = 347.63) compared to the reduced model (*AIC* = 345.56, *BIC* = 357.06). Additionally, the bayes factor is 0.009 for the reduced model, and 111.43 for the full model (over the 100+ metric) signifying decisive evidence in favor of the full model. The full model explains about 8.4% more variance in negative affect than sex and conscientiousness alone (*Table 2*). This means that scores on SPP are clinically meaningful for predicting negative affect above and beyond sex and conscientiousness scores.

We can conclude that H1 is partially supported. Results showed that conscientiousness is negatively related to negative affect and higher SPP is associated with higher negative affect. However, we did not see a significant associated between sex and negative affect. H2 was supported as SPP did predict negative affect over and above conscientiousness and sex in a meaningful way. It should be noted that because our assumptions are violated, we cannot trust our results. There may be a better model to represent this data out there (perhaps a curve) so future studies should compare this linear model with other model variations.

**References**

Philipp, M. (2021). flexplot: Graphics package to enhance the ggplot2 plotting system. Retrieved from <https://cran.r-project.org/package=flexplot>

Stanley, D. (2021). apaTables: Create American Psychological Association (APA) style tables. Retrieved from <https://cran.r-project.org/package=apaTables>

Venables, W. N., & Ripley, B. D. (2002). Modern applied statistics with S (4th ed.). Springer.

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T. L., Miller, E., Bache, S. M., Müller, K., Ooms, J., Robinson, D., Seidel, D. P., Spinu, V., Takahashi, K., Vaughan, D., Wilke, C., Woo, K., & Yutani, H. (2019). Welcome to the tidyverse. *Journal of Open Source Software, 4*(43), 1686. <https://doi.org/10.21105/joss.01686>

**Appendices**

**A screenshot of a white sheet

Description automatically generated**

A screenshot of a report

Description automatically generated