**Assignment 2: General Linear Modelling**

Colby Price

Department of Psychiatry; Dalhousie University

PSY-6003: Fundamentals of Applied Statistics and Research Design

Dr. Igor Yakovenko

April 6, 2024

**Assignment 2: General Linear Modelling**

In this study, undergraduate students (*n*=137) were gathered to complete three subscales for their psychology statistics course, which included the 3-year versions of the Multidimensional Perfectionism Scale Short Form: Socially Prescribed Subscale (MPS; Hewitt & Flett, 2004), the Positive and Negative Affect Scales (PANAS; Watson et al., 1988), and the Ten Item Personality Measure: Conscientiousness Subscale (TIPI; Gosling et al., 2003). Participants were given a responder ID and filled out the following scales and provided their biological sex. In total, 21 variables were recorded for each participant, and mean scores were tabulated for each subscale. In our sample, one participant was removed from our data due to being intersex (*n*=136). As only one participant belonged to the intersex category, our team removed them as they would not provide generalizable or useful data for the purposes of our study. Five cases were removed from our dataset for incomplete answers (*n*=131).

Our study objective was to analyze the relationship of sex, conscientiousness, and socially prescribed perfectionism (SPP) on negative impact. Our study’s hypotheses were 1) sex, conscientiousness, and SPP will predict negative affect and 2) SPP will predict negative affect over and above sex and conscientiousness. All data management and statistical analyses were conducted via R version 4.3.2 (2023) using “tidyverse” (Wickham et al., 2019), “flexplot” (Fife, 2024), “haven” (Wickham et al., 2023), “MASS” (Venables & Ripley, 2002), “patchwork” (Presden, 2024), and “apaTables” packages (Stanley, 2021).

**Data Analytic Plan**

A general linear model comparison with two numeric variables and one categorical variable will be used to establish general trends within our data. Our research team will visualize several different model types (e.g. linear and quadratic) to find the model of best fit according to our study’s objectives. A sensitivity analysis will be run to determine if our models violate assumptions of normality, dependence, and homoskedasticity; consequently, if any assumptions are violated, a robust errors model will be created to attempt to correct said violations. If these assumptions are still violated, our team will continue the analysis using the original model of best fit.

A model comparison test will be done between a full model, which will include our study’s objective of SPP, and a reduced model, which will exclude the SPP measure. After accepting either model, estimates, standard deviations, and semi-partial *r2* will be collected for each parameter of the respective model. Furthermore, the model’s *R*2 will also be reported.

**Descriptive Statistics and Bivariate Correlations**

Our sample was largely comprised of female participants (*n=*112) with the remaining participants identifying as male. Means, standard deviations, and correlation coefficients of total objective measurement scores and their interactions are described in Table 1. The lowest mean scores were gathered by the PANAS with average negative affect scores being lower (*M* =2.44 (*SD* =0.97)) than both the MPS and TIPI scores. MPS mean scores were *M=*4.38 (*SD =*1.29), while TIPI mean scores were *M* = 5.03 (*SD* =1.43).

In the correlation analysis, we have denoted notable weak interactions between the TIPI, MPS, and the PANAS. Weak negative relationships exist between conscientiousness and negative affect (*r* = -0.37) and between conscientiousness and SPP (*r* =-.21), meaning that as conscientiousness decreases, negative affect and SPP increase. Moreover, a notable positive relationship exists between the SPP and negative affect variables (*r* =.37), meaning that SPP increases with negative affect scores.

**Results**

After visualizing our models, our team determined the quadratic model best fit our data. Sensitivity analyses showed that, while our model was normally distributed, the assumptions of dependence and homoskedasticity were violated. After visualizing a robust error model of our quadratic equation, these assumptions were still violated; therefore, we continued the analysis using our original quadratic model.

Our team chose a full model as compared to a reduced model via a model comparison test using the ‘flexplot’ package (Fife, 2024). The model comparison test showed that the full model was a better fit to our data set, as indicated by both higher *R2*(*R2=*.299) and a greater Baye’s factor (*BF*=31.36, meaning strong likelihood). Moreover, both AIC and BIC scores were lower for the full model, and the p-values indicated significant differences between the models (*p*=2.0x1016). Coefficients and similar results for each variable are tabulated in Table 2. Our model accounts for 29.9% of negative affect (as indicated by our *R2*). SPP predicts unique variance as compared to the predictors of conscientiousness and sex, as shown by a greater semi-partial *r2* (*rMPS2=*0.134 and *rMPS22=*0.018; *r2=*0.152). This data would mean that SPP accounts for 15.2% of negative affect scores within our model.

In the case of our model, for every one-point increase in SPP within our model, negative affect would decrease by *b=*0.22 (*bMPS*=-0.28and *bMPS2*=0.06); furthermore, a point increase in TIPI, would decrease negative affect by *b=­*0.69 (*bTIPI*=-0.74 and *bTIPI2*=0.05). Sex was a relevant predictor, with a difference score of *bSex=*-0.51 between male and female scores, meaning females had scores .51 points higher as compared to males. As for our reported intercept value, when all predictor variables are equal to zero in our model, negative affect is *b0=*4.92. In conclusion, the following model supports both our hypotheses that SPP, conscientiousness, and sex are relevant predictors for negative affect scores according to the PANAS and that SPP predicts a unique variance as compared to other predictor variables.

**References**

Fife D (2024). \_flexplot: Graphically Based Data Analysis Using 'flexplot'\_. R package version 0.20.3.

Gosling, S. D., Rentfrow, P. J., & Swann, W. B., Jr. (2003). A Very Brief Measure of the Big Five Personality Domains. Journal of Research in Personality, 37, 504-528.

Hewitt, P. L., & Flett, G. L. (2004). [Multidimensional Perfectionism Scale (MPS): Technical manual.](https://storefront.mhs.com/collections/mps)Toronto, Canada: Multi-Health Systems.

Pedersen T (2024). \_patchwork: The Composer of Plots\_. R package version 1.2.0,

<https://CRAN.R-project.org/package=patchwork>.

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>.

Venables, W. N. & Ripley, B. D. (2002) Modern Applied Statistics with S. Fourth Edition. Springer, New York. ISBN 0-387-95457-0

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: the PANAS scales. Journal of personality and social psychology, 54(6), 1063.

Wickham H, Averick M, Bryan J, Chang W, McGowan LD, François R, Grolemund G, Hayes A, Henry L, Hester J, Kuhn M, Pedersen TL, Miller E, Bache SM, Müller K, Ooms J, Robinson D, Seidel DP, 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. doi:10.21105/joss.01686 <https://doi.org/10.21105/joss.01686>

Wickham H, Miller E, Smith D (2023). \_haven: Import and Export 'SPSS', 'Stata' and 'SAS' Files\_. R package version 2.5.4, <https://CRAN.R-project.org/package=haven>.

**Table 1**

*Descriptive statistics and correlations sex, TIPI, PANAS, and MPS scores*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | *M* (*SD*) | Sex | MPS | PANAS | TIPM |
| Sex | 1.15 (0.35) | --- |  |  |  |
| MPS | 5.03 (1.29) | 0.01  (-0.16, 0.18) | --- |  |  |
| PANAS | 2.44 (0.98) | -0.16  (-0.32, 0.01) | 0.37  (0.21, 0.51) | --- |  |
| TIPI | 4.39 (1.43) | -0.14  (-0.30, 0.03) | -0.21  (-0.37, -0.04) | -0.37  (-0.51, -0.22) | --- |

**Table 2**

*General linear model estimates, standard error, 95% confidence intervals, semi-partial R2, and R2*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | *E* (*SE*) | 95% Confidence Interval | *Semi-Partial R2* | *R2* |
| Intercept | 4.92 (0.99) | (2.99, 6.85) | --- | --- |
| MPS Mean | -0.28 (0.27) | (-0.81, 0.26) | 0.134 | --- |
| TIPI Mean | -0.74 (0.34) | (-1.40, -0.09) | 0.092 | --- |
| Sex (Male) | -0.51 (0.21) | (-0.93, -0.09) | 0.043 | --- |
| (MPS Mean)2 | 0.06 (0.03) | (-0.01, 0.12) | 0.018 | --- |
| (TIPI Mean)2 | 0.05 (0.03) | (-0.02, 0.12) | 0.012 | --- |
| --- | --- | --- | --- | 0.299 |