**Assignment 2: The General Linear Model**

Mikaela Ethier-Gagnon

Department of Psychiatry, Dalhousie University

PSYR6003: Fundamentals of Applied Statistics

Dr. Igor Yakovenko

April 6th, 2024

**Assignment 2: The General Linear Model**

Prior literature suggests that sex, conscientiousness, and socially prescribed perfectionism are associated with negative affect. In the current investigation, we hypothesized that sex, conscientiousness, and socially prescribed perfectionism would all significantly predict negative affect. Specifically, women were anticipated to demonstrate increased negative affect relative to men, conscientiousness was expected to be negatively related to negative affect, whereas socially prescribed perfectionism was anticipated to be positively associated. Moreover, given that sex, conscientiousness, and socially prescribed perfectionism all tend to covary, we hypothesized that socially prescribed perfectionism would predict unique variance in negative affect over and above sex and conscientiousness in a meaningful way. In the present study, data was collected from undergraduates (*N*=137). Participants were assessed for demographic characteristics (i.e., sex), socially prescribed perfectionism (SPP) using the Multidimensional Perfectionism Scale Short Form: Socially Prescribed Subscale (3 year version), negative affect with the PANAS: Negative Affect (3 year version) scale, and conscientiousness using the Ten Item Personality Measure: Conscientiousness Subscale (3-year version). This dataset was obtained from Dr. Igor Yakovenko’s repository on GitHub <https://github.com/iyakoven/PSYR6003-Assignment-2>. Data analysis was completed using a custom R script in R Studio (version 4.3.2; R Core Team, 2023). Incomplete cases and participants identifying as non-binary were excluded from the dataset, resulting in a final sample of *N*=132.

Descriptive statistics for included participants were calculated using the *apaTables* package (Stanley, 2021). Outcomes of interest included SPP (*M=*4.40, *SD*=1.44), conscientiousness (*M=*5.02, *SD*=1.29), negative affect (*M=*2.43, *SD*=0.97), and sex (16.8% male). Negative affect was weakly positively correlated with SPP (*r*=0.36, *p* < .01), and negatively correlated with conscientiousness (*r*=-0.37, *p* < .01). Conscientiousness and SPP were also demonstrated to covary as they were weakly positively correlated (*r*=0.19, *p* < .05). Contrary to prior literature, sex was not found to be significantly correlated with negative affect (*r*=-0.16, *p* > .05), SPP (*r*=0.01, *p* > .05), or conscientiousness (*r*=-0.13, *p* > .05). Total scores for SPP, conscientiousness, and negative affect represented summed means of the items within their respective scales. Total scores were generated using the *Tidyverse* package (Wickham et al., 2019). Means, standard deviations, and bivariate correlations are presented in Table 1.

We aimed to determine the impact of SPP, conscientiousness, and sex on self-reported negative affect using a multiple regression model. Upon visualization of the univariate distributions, there appeared to be more women than men in the sample, both conscientiousness and SPP were positively skewed, and negative affect scores were negatively skewed. Statistical assumptions of the model including normality, linearity, homoscedasticity, and independence of observations were tested using the *Flexplot* package (Fife, 2024). Specifically, to test normality of residuals we derived a histogram of residuals. The observed pattern was relatively normally distributed with a slight negative skew. To determine whether this slight violation of normality greatly impacts the model, we conducted a sensitivity analysis using the *MASS* package (Venables et al., 2002). We fit a linear model by robust regression and the line of best fit did not appear to differ from the original model, so the violation of normality was deemed unimportant for the fit of the current model. Moreover, there appeared to be a violation of linearity as evidenced by a slightly curved line in the residual dependence plot. Therefore, we added a quadratic term to the linear model to account for non-linearity. The residual dependence plot was more linear than when previously fitted to the linear model, suggesting that the linear model fit was superior to the quadratic. Finally, the assumption of homoscedasticity was also 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. Given that we were not assessing repeated measures or participants categorized into groups, the assumption of independence of observations was met due to experimental design. Estimates and 95% Confidence Intervals of the model are presented in Table 2. Consistent with our hypothesis, SPP, conscientiousness, and sex were found to be predictive of negative affect *R*2 = .27, *F*(3, 128) = 15.79, CI [0.14, 0.37]. The full model explained 27% of the variance in negative affect. Notably, conscientiousness accounted for most of the effect; specifically; 15.8% of the variance in negative affect, whereas sex accounted for 2.6%, and SPP accounted for 8.6% of the variance. If participants were female and had a score of 0 for SPP and conscientiousness, they would have a negative affect score of 2.91. For every one-point increase in conscientiousness, there was a 0.26-point decrease in negative affect. Similarly, individuals identifying as male had a -0.57-point decrease in negative affect. Finally, for every one-point increase in SPP, there was a 0.20-point increase in negative affect. The full model could potentially predict up to a 2.26-point differential in negative affect. Thus, as hypothesized, women tended to have more negative affect than men, increased levels of conscientiousness were associated with decreased levels of negative affect, and higher SPP was associated with higher negative affect.

We also aimed to examine whether SPP predicted unique variance in negative affect above and beyond sex and conscientiousness in a meaningful way. Using the *Flexplot* package (Fife, 2024), we conducted model comparison to determine whether the fit of the full model was improved when including SPP relative to the reduced model including only conscientiousness and sex to justify the added complexity. Upon visual inspection, there appeared to be a very weak positive relationship between SPP and negative affect when controlling for conscientiousness and sex. Consistent with our hypothesis, fit indices suggested that SPP was predictive of negative affect while controlling for sex and conscientiousness. Notably, the full model explained 8.6% more variance in negative affect and resulted in a 0.65-point differential in prediction relative to the reduced model. Also, the bayes factor was >100 for the full model, suggesting that there was decisive evidence to support use of the full model. Moreover, The Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) were both lower in reference to the full model as compared to the reduced model, suggesting that the full model was a better fit. Finally, the full model was found to be significantly different from the reduced model (*p*<0.001). Therefore, consistent with our hypothesis, we can conclude that SPP does predict negative affect above and beyond sex and conscientiousness.

**References**

Cumming, G. (2014). The New Statistics: Why and How. Psychological Science, 25(1), 7-29. https://doi-org.ezproxy.library.dal.ca/10.1177/0956797613504966

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>.

Venables, W. N., & Ripley, B. D. (2002). \_Modern Applied Statistics with S.\_ (4th ed.). Springer, New York. ISBN 0-387-95457-0.

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

Table 1

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

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Variable | *M* | *SD* | 1 | 2 | 3 |
|  |  |  |  |  |  |
| 1. SPP | 4.38 | 1.42 |  |  |  |
|  |  |  |  |  |  |
| 2. Conscientiousness | 5.01 | 1.31 | -.19\* |  |  |
|  |  |  | [-.35, -.02] |  |  |
|  |  |  |  |  |  |
| 3. Negative Affect | 2.44 | 0.97 | .36\*\* | -.37\*\* |  |
|  |  |  | [.20, .50] | [-.51, -.22] |  |
|  |  |  |  |  |  |
| 4. Sex | 0.14 | 0.35 | .01 | -.13 | -.16 |
|  |  |  | [-.16, .18] | [-.29, .04] | [-.32, .01] |
|  |  |  |  |  |  |

*Note.* *M* and *SD* are used to represent mean and standard deviation, respectively. SPP represents socially prescribed perfectionism. Sex was recoded to a factor; 0=Female, 1=Male. 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

*Regression results using negative affect as the criterion*

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Predictor | *b* | *b*  95% CI  [LL, UL] | *beta* | *beta*  95% CI  [LL, UL] | *SE* | *sr2* | *sr2*  95% CI  [LL, UL] | *r* | Fit |
| (Intercept) | 2.91\*\* | [2.10, 3.73] |  |  | 0.41 |  |  |  |  |
| Sex | -0.57\*\* | [-0.99, -0.16] | -0.21 | [-0.36, -0.06] | 0.21 | .04 | [-.02, .10] | -.16 |  |
| Conscientiousness | -0.26\*\* | [-0.37, -0.14] | -0.34 | [-0.50, -0.19] | 0.06 | .11 | [.02, .21] | -.37\*\* |  |
| SPP | 0.20\*\* | [0.10, 0.31] | 0.30 | [0.15, 0.45] | 0.05 | .09 | [.00, .17] | .36\*\* |  |
|  |  |  |  |  |  |  |  |  | *R2*  = .270\*\* |
|  |  |  |  |  |  |  |  |  | 95% CI[.14,.37] |
|  |  |  |  |  |  |  |  |  |  |

*Note.* SPP represents socially prescribed perfectionism. A significant *b*-weight indicates the beta-weight and semi-partial correlation are also significant. *b* represents unstandardized regression weights. *beta* indicates the standardized regression weights. *sr2* represents the semi-partial correlation squared. *r* represents the zero-order correlation. *LL* and *UL* indicate the lower and upper limits of a confidence interval, respectively.  
\* indicates *p* < .05. \*\* indicates *p* < .01