**PSYR 6003 Assignment 1: Database Management and Sample Size Estimation**

Holly Carbyn

Department of Psychiatry, Dalhousie University

PSYR 6003: Fundamentals of Applied Statistics

Dr. Igor Yakovenko

March 7th, 2024

**Assignment 1: Database Management and Sample Size Estimation**

In this data analysis, we evaluated the fictional “Avengers” dataset (*n* = 814) using statistical software R version 4.3.2 (Eye Holes). This dataset was gathered as part of the Avengers’ final battle against Thanos across two battlefields (North and South). It includes variables such as “IQ” (the IQ of the avenger), “Superpower” (describes whether the avenger has superpowers or not), “Died” (did the avenger die or not), etc. I used the software “Github” to clone the repository from Dr. Igor Yakovenko. We used the ‘Tidyverse’ package (Wickham et al., 2019) to conduct analyses as the purpose of this assignment was to demonstrate our ability to manage data using the ‘Tidyverse’ packages. The dataset was cleaned by removing two cases with missing values (*n* = 812).

**Descriptive Statistics**

Descriptive statistics were calculated for the avengers who died and who did not have superpowers. Our variables of interest included combat effectiveness; which is the sum of agility, speed, strength, and willpower (*M* = 497.54, *SD* = 177.56, range = 67.25–946.89), how many injuries each avenger sustained (*M* = 4.55, *SD* = 0.74, range = 2-5), and how many bad guys they each killed (*M* = 2.55, *SD* = 8.81, range = 0-79). It is important to note that the error in the mean model for combat effectiveness was relatively high (*SD* = 177.56) compared to number of injuries and number of kills (*SD* = 0.74 and *SD* = 8.81). This means that the variable combat effectiveness is likely the most erroneous in our model and a mean model might not be the best descriptive tool for it.

Results showed that the Northern battlefield was more effective in combat (*M* = 499.78, *SD* = 174.07, range = 766.38) than the Southern battlefield (*M* = 491.68, *SD* = 189.53, range = 879.64). However, the Northern battlefield also had more injuries (*M* = 4.60, *SD* = 0.68, range = 2-5) than the Southern battlefield (*M* = 4.43, *SD* = 0.88, range = 2-5).

**Power Analysis**

In the secondary analysis, we aimed to understand the relationship between having superpowers and IQ for avengers. We hypothesized that avengers with superpowers would have a higher average IQ than those without.

To estimate the required sample for this study, we ran an a-prior power analysis using the ‘pwr’ package (Champely, 2020). We ran an a-prior power analysis to estimate sample size because our research question aims to test whether the effect size can be statistically rejected with a desired statistical power. Another alternative would be to measure the entire population of avengers because this would likely be accessible due to the limited number of avengers and the finite specific nature of the population. The effect size we are looking for (*d* = 0.3) as this is our smallest effect size of interest (SESOI). *d =* 0.3 is the smallest effect size we could get that would be clinically/practically relevant. A smaller effect size would not be enough to justify putting the money into creating a new gifted school for children with superpowers.

We carried out a power analysis for a two-sided independent samples t-test because there is limited literature on the relationship between IQ and superpowers. Although we hypothesized that avengers with superpowers would have a higher average IQ than those without, there was not enough evidence to support the direction of this relationship. I chose to use power of 0.8 for this analysis because we are on a very tight budget for this study, and it is often deemed the lowest acceptable unit for power. Additionally, previous literature on this topic uses power of 0.8. We are using significance level of α = 0.05 as we are only running one statistical test. With this significance level there is a 5% chance that the observed result is due to chance. Our power analysis showed that we would need a sample size of 176 people per group to run the proposed study (*n* = 175.38) comparing avengers with and without superpowers on IQ.

**Equivalence Test**

To ensure that we would have enough power to confirm if there truly is no difference between the groups (zero effect), we conducted an equivalence test. This requires specifying an equivalence range of values that we do not care about (-0.3 to 0.3), which equates a zero effect for our proposed study. Using the ‘TOSTER’ package (Caldwell, 2022; Lakens, 2017) an equivalence test was conducted using a significance level of α = 0.05, lower and upper bounds of -0.3 and 0.3, and our previously calculated sample size of 176 per group. The estimated power required to ensure a zero effect is 0.76 as determined by the equivalence analysis.

**Effect Size and Confidence Interval**

We conducted the independent t-test analysis to find the test statistic (*t* = 4.25). The ‘effectsize’ package (Ben-Shachar, Lüdecke, & Makowski, 2020) was used to calculate a relevant effect size (*d* = 0.03) and 95% confidence interval (CI) for the effect [0.16, 0.44]. This 95% CI unfortunately is not a precise enough estimate in for our study. Although this estimate contains the calculated small effect of *d* = 0.3, the lower bound is less than our SESOI. As previously stated, we do not care about an effect that is less than *d* = 0.3, so a lower bound of 0.16 indicates that the effect might be too small.

**References**

Ben-Shachar, M., Lüdecke, D., & Makowski, D. (2020). effectsize: Estimation of Effect Size Indices and Standardized Parameters. *Journal of Open Source Software, 5*(56), 2815. <https://doi.org/10.21105/joss.02815>

Caldwell, A. R. (2022). Exploring Equivalence Testing with the Updated TOSTER R Package. *PsyArXiv*. <https://doi.org/10.31234/osf.io/ty8de>

Champely, S. (2020). *pwr: Basic Functions for Power Analysis* (Version 1.3-0). Retrieved from <https://CRAN.R-project.org/package=pwr>

Lakens, D. (2017). Equivalence tests: A practical primer for t-tests, correlations, and meta-analyses. *Social Psychological and Personality Science, 8*(4), 355–362. <https://doi.org/10.1177/1948550617697177>

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>

Lakens, D. (2017). Equivalence tests: A practical primer for t-tests, correlations, and meta-analyses. *Social Psychological and Personality Science, 8*(4), 355–362. https://doi.org/10.1177/1948550617697177