Review of: "Intersections of Statistical and Substantive Significance Under a True and False Null Hypothesis"

Heitor Marques Honório[1]

1 Universidade de São Paulo

Potential competing interests: No potential competing interests to declare.

The article "Intersections of Statistical and Substantive Significance Under a True and False Null Hypothesis" by Eugene Komaroff presents a thorough and insightful examination of the interplay between statistical and substantive significance. The use of histograms, bar graphs, and crosstabs to elucidate these concepts is commendable, providing readers with a clear and intuitive understanding of how p-values and effect sizes can be interpreted. The empirical approach using simulated data sets is particularly effective in demonstrating the practical implications of statistical significance, especially in the context of small sample sizes.

One of the strengths of this study lies in its methodological rigor. The simulations are well-designed, and the decision to include various sample sizes (n = 15, 64, 500, 1000) allows for a comprehensive analysis of how statistical significance behaves under different conditions.

However, there are areas where the paper could be improved. One critique is that the discussion on the misuse and misunderstanding of p-values, while important, could be expanded to include more recent debates in the statistical community. For instance, the paper mentions the American Statistical Association's (ASA) statement on p-values but does not delve deeply into the various viewpoints that have emerged since then. Including a more nuanced discussion on the ongoing debates around the use of p-values and alternative approaches, such as Bayesian methods or the use of confidence intervals, would add depth to the analysis.

Another area for improvement is the presentation of the results. While the graphs and tables are informative, the narrative could benefit from a more detailed interpretation of the findings. For example, the discussion on the implications of the different sample sizes could be enriched by linking the results more explicitly to practical scenarios that researchers might face. This would help readers better understand how to apply these findings in their own work.

Furthermore, the conclusion section, while summarizing the key points effectively, could be strengthened by providing more concrete recommendations for researchers. Rather than simply stating that statistical significance is useful with small sample sizes, the paper could offer guidelines on how to balance statistical and substantive significance in practice. This might include advice on how to report results in a way that clearly distinguishes between statistical significance and practical relevance.

Lastly, while the paper is well-written, there are a few areas where the language could be more precise. For instance, the term "substantive significance" is used throughout the paper, but its definition could be clarified early on to ensure that all readers are on the same page.

In conclusion, Eugene Komaroff's paper is a valuable contribution to the literature on statistical significance, offering important insights and practical guidance. By expanding the discussion on current debates, providing more detailed interpretations of the results, and offering concrete recommendations, the paper could be further strengthened and have an even greater impact on its readers.