Review of: "Intersections of Statistical Significance and Substantive Significance: Pearson's Correlation Coefficients Under a Known True Null Hypothesis"

Nabil Semmar

Potential competing interests: No potential competing interests to declare.

Taking into account the fact that the manuscript would be intended for non-statisticians, some structuration of presentation would be required:

1. The introduction did not talk about the content of the manuscript to prepare (non-statistician) readers for easy understanding of the presented work. For instance, the correlation concept was not evoked in the introduction, whereas it represents the backbone of the manuscript. The current version shows jumping between the introduction and the manuscript content (and Discussion).

2. Also, the paragraphs of the introduction should be structured through sequentially or gradually linked ideas associated with the whole content of the manuscript (problematic, methods, general aspects of results, usefulness).

3. I previously recommended a synthesis aspect of figures, but no revision or effort was made in this way. I invite the author to see that Figures 5, 10, 15, 20, 25 are redundant with Figures 4, 9, 14, 19, 24, respectively. Therefore, the presentation of Figures 4, 9, 14, 19, 24 would be sufficient.

4. At the end of the manuscript, a synthesis figure on the whole simulation results would be recommended. The style of writing and presentation doesn't help to deduce the ultimate goal of the manuscript. The manuscript contains many figures and tables presented under a cumulated and cyclic way. The current discussion doesn't return to the results to discuss them in relation with limits, advantages, and perspectives. This doesn't help non-statisticians for the contemplated intuitive understanding.

To help non-statisticians benefit from an integrative picture and discussion of the work, I suggest an integrative figure issued from the synthesis of the whole work that could be used as a basis for the Discussion section (**see attached file**). For instance:

- The SE converges under sample size n=1000. However, more sampling (additional sampling) would be required between n=100 and 1000 to determine more precisely the sampling size of convergence.

- However, the percentage of significance converges from sample size n=30.

- The effect size needs a sampling size higher than n=1000 (2000) to be eliminated; Therefore, additional sampling between 1000 and 2000 usefully helps for a more precise determination of the minimal required sampling size.

- Finally, the number of significant correlations seems to be ready for stationarity from n=30, but the applied sampling design (n=4, 30, 100, 1000, 2000) did not seem to be sufficient to cover or to highlight stationarity. However, the min-

max of r reached a precise and stationary variation range from sample size n = 1000.

These synthesis results show different sensitivities or behaviors of different parameters toward sample sizes. This suggestion of a synthetic figure could help the author to improve the readability and understanding of the current manuscript content by non-statisticians.

Minor comments:

5. The revised version was quasi-similar to the previous one. Some added legends should be improved: for instance, |r|=0.4 belongs both in the ranges |r|≥0.10 and |r|≥0.30. Typos concerning inequality signs should be corrected to allow clear classification of r values.

6. The "r" in the parameter "zr" should be in index. This helps non-statisticians to easily follow the manuscript content.