Review of: "A New Index for Measuring the Difference Between Two Probability Distributions"

Sooyong Lee[1]

1 University of Wisconsin-Madison

Potential competing interests: No potential competing interests to declare.

Review of the Paper on the Distribution Discrepancy Index (DDI)

This paper introduces the Distribution Discrepancy Index (DDI), a novel metric designed to measure the differences between two probability distributions. Based on the newly proposed informity theory, the DDI utilizes the concepts of informity, cross-informity, and informity divergence. The DDI is quantitatively defined as the ratio of informity divergence between two distributions to the sum of their respective informities. With values ranging from 0 to 1, the DDI provides an intuitive scale where lower values indicate minimal differences and higher values suggest significant disparities between distributions.

Questions and Suggestions

Q1. Interpretation of DDI Values

- The paper provides benchmarks for the DDI, with values of 0.25, 0.5, and 0.75, suggesting small, moderate, and high levels of difference, respectively.
- While these thresholds are helpful, the paper should clarify whether the relationship between DDI values and the degree of difference is linear, as figures in the paper show non-linearity.
- Could you elaborate on the rationale behind these specific thresholds, or at least mention that the interpretation of DDI values should be done with caution due to the non-linear relationship?

Q2. Practical Applications of DDI

- The paper could greatly benefit from examples of how and when (and in what circumstances) the DDI can be applied in practical settings.
- For instance, the DDI might be useful in evaluating classification quality within a mixture modeling framework or in selecting items in computerized adaptive testing, where Kullback-Leibler divergence is currently employed. These are just a few examples that come to mind, but the author may have better ideas about the practical applications of DDI.
- Including such examples of applications could provide valuable insights for applied researchers.

Q3. Addressing Non-Normal Distributions

- Many real-world continuous variables, such as Body Mass Index (BMI), do not follow a normal distribution and may exhibit skewness.
- It would be beneficial to say something like “future research directions could address the application of DDI to non-normal distributions.”

Q4. Impact of Varying Distribution Parameters

- The study investigates how changes in the means and variances of distributions affect the DDI. However, it would be insightful to explore situations where both parameters change simultaneously.
- Additionally, a discussion on which parameter (mean or variance) more significantly influences DDI changes would enrich the paper’s analysis.

Q5. Comparison with Other Distance Measures

- While the paper mentions several established distance measures, it would strengthen the argument for using DDI by clearly contrasting it with alternatives.
- What specific advantages does DDI offer over these methods in terms of accuracy, interpretability, or simplicity?
- This comparison is crucial for justifying the adoption of DDI in research and application.

By addressing these queries and suggestions, the paper could significantly enhance its contributions to the field of statistical analysis and its utility in applied research contexts.