

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

Jose Oliveira Siqueira¹

¹ Universidade de São Paulo

Potential competing interests: No potential competing interests to declare.

Peer Review Report

Title: A New Index for Measuring the Difference Between Two Probability Distributions

Author: Hening Huang

General Comments

The paper presents a novel and promising index for measuring the difference between two probability distributions, named the Distribution Discrepancy Index (DDI). While the concept introduced is innovative, the paper is notably concise. To enhance the impact and clarity of the proposed measure, several key aspects need further elaboration.

Specific Comments

- 1. Lack of Detailed Properties:** The paper would benefit significantly from a detailed exploration of the properties of the new measure. A thorough analysis of its behavior and characteristics under various conditions is essential to understand its robustness and applicability fully.
- 2. Additional Distribution Examples:** Currently, the paper provides limited examples to demonstrate the application of the DDI. To illustrate the versatility and reliability of the new measure, it would be advantageous to include comparisons across a wider range of distributions. Suggested distributions include lognormal, triangular, beta, gamma, binomial, Poisson, and geometric distributions. This expansion would help validate the measure across different statistical scenarios.
- 3. Proof of DDI Range (0 to 1):** The paper should explicitly demonstrate that the DDI ranges between 0 and 1. While this is stated, a mathematical proof or detailed explanation of why this range holds would add rigor to the presentation and help readers understand the bounds of the measure.
- 4. Application Examples in R or Python:** For the practical application of the proposed index, it would be beneficial to include examples solved using R or Python. Providing the scripts for these examples would not only enhance reproducibility but also aid readers in applying the DDI in their own work. This practical component would make the paper more accessible to a broader audience, including those with varying levels of expertise in statistical programming.

5. PSI Formula in Example 1: In the first example regarding the grade distribution of credit scores, it is essential to include and explain the formula for the Population Stability Index (PSI). This will provide readers with a complete understanding of the comparison being made and how the DDI relates to other established indices.

Conclusion

The proposed Distribution Discrepancy Index (DDI) is a potentially valuable tool for comparing probability distributions. However, the current manuscript needs significant expansion and clarification to fully convey its utility and robustness. By addressing the points mentioned above, the paper will provide a more comprehensive and convincing presentation of the DDI.

I recommend revising the manuscript to incorporate these suggestions before considering it for publication.