

Work Instruction: Choosing Between Cohen's Kappa and Fleiss' Kappa in Attribute Agreement Analysis

Purpose

This work instruction describes how to choose between Cohen's Kappa and Fleiss' Kappa in attribute agreement analysis. The purpose of this document is to help analysts determine which method is appropriate for their specific situation.

Scope

This work instruction applies to all attribute agreement analyses conducted in the organization.

Definitions

- Attribute Agreement Analysis: A method used to evaluate the consistency and reliability of a measurement system by comparing the results of multiple operators measuring the same characteristic on the same part.
- Cohen's Kappa: A statistical measure of inter-rater agreement for categorical items. It is used to assess the degree of agreement between two raters beyond chance.
- Fleiss' Kappa: A statistical measure of inter-rater agreement for categorical items. It is used to assess the degree of agreement between multiple raters beyond chance.

Procedure

1. Determine the number of raters involved in the study. If there are only two raters, use Cohen's Kappa. If there are three or more raters, use Fleiss' Kappa.
2. Determine the type of agreement being measured. If the agreement is nominal, use Cohen's Kappa. If the agreement is ordinal or nominal, use Fleiss' Kappa.
3. Calculate the appropriate kappa coefficient using the chosen method.
4. Interpret the kappa coefficient according to the guidelines provided in the literature or the relevant standards.
5. Document the method used and the interpretation of the results in the attribute agreement analysis report.

References

- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20, 37–46.
- Fleiss, J. L. (1971). Measuring nominal scale agreement among many raters. *Psychological Bulletin*, 76(5), 378–382.
- ISO 5725-2:1994. Accuracy (trueness and precision) of measurement methods and results - Part 2: Basic method for the determination of repeatability and reproducibility of a standard measurement method.
- AIAG MSA 4th Edition, 2010.