Welcome to the Kappa Zoo

David Eubanks

Table of contents

1 Welcome

This site explains rater agreement and associated reliability statistics. Much of the content is new, using a general model to explain the features of existing methods from the measurement literature, in the context of more modern approaches from the machine learning (ML) literature. The synthesis of these two fields is the t-a-p model, which combines the intuitiveness of the measurement ideas with the statistical power of the ML algorithms.

If you are new to t-a-p models, I recommend starting with the Chapter 1, which introduces rater agreement as a classification problem and lays out the philosophical and statistical assumptions. Chapter 2 provides the basic statistical derivations needed to work with t-a-p models. Chapter 3 derives the relationship to some existing rater agreement statistics, for example showing how the Fleiss kappa is a special case of a t-a-p model. Chapter 4 expands the number of parameters and shows how this relates to the "kappa paradox." Chapter 5 allows each rater and/or subject to have individual parameters for accuracy and truth, and allows for independent variables to be used as regression inputs. Chapter 6 reviews some of the machine learning algorithms similar to t-a-p models. Chapter 7 introduces the `tapModel` R package, which allows for easy estimation of the model parameters. Chapter 8 introduces the t-a-p app, an interactive way to use real or simulated data to estimate the model parameters without coding.

An appendix is included to provide more detailed statistical derivations and proofs for some results.

The Kappa Zoo will eventually include a collection of real-world data sets and the model parameters estimated from them.

This is a work in progress. The unlinked chapters aren't finished yet.

1.1 Table of Contents