



Evaluation of probabilistic forecasts with the `scoringRules` package

Alexander Jordan (1), Fabian Krüger (2), Sebastian Lerch (1,3)

(1) Heidelberg Institute for Theoretical Studies, Heidelberg, Germany, (2) Heidelberg University, Heidelberg, Germany, (3) Karlsruhe Institute of Technology, Institute for Stochastics, Karlsruhe, Germany (sebastian.lerch@h-its.org)

Over the last decades probabilistic forecasts in the form of predictive distributions have become popular in many scientific disciplines. With the proliferation of probabilistic models arises the need for decision-theoretically principled tools to evaluate the appropriateness of models and forecasts in a generalized way in order to better understand sources of prediction errors and to improve the models.

Proper scoring rules are functions $S(F, y)$ which evaluate the accuracy of a forecast distribution F , given that an outcome y was observed. In coherence with decision-theoretical principles they allow to compare alternative models, a crucial ability given the variety of theories, data sources and statistical specifications that is available in many situations.

This contribution presents the software package `scoringRules` for the statistical programming language R, which provides functions to compute popular scoring rules such as the continuous ranked probability score for a variety of distributions F that come up in applied work. For univariate variables, two main classes are parametric distributions like normal, t , or gamma distributions, and distributions that are not known analytically, but are indirectly described through a sample of simulation draws. For example, ensemble weather forecasts take this form.

The `scoringRules` package aims to be a convenient dictionary-like reference for computing scoring rules. We offer state of the art implementations of several known (but not routinely applied) formulas, and implement closed-form expressions that were previously unavailable. Whenever more than one implementation variant exists, we offer statistically principled default choices. Recent developments include the addition of scoring rules to evaluate multivariate forecast distributions. The use of the `scoringRules` package is illustrated in an example on post-processing ensemble forecasts of temperature.