Universal Inference for Testing Calibration of Mean Estimates within the Exponential Dispersion Family

Łukasz Delong [*1] and Mario Wüthrich [†2]

[1]University of Warsaw, Department of Statistics and Econometrics
[2]ETH Zürich, Department of Mathematics

Abstract

Calibration of mean estimates for prediction is a crucial property in many applications, particularly in the fields of financial and actuarial decision making. In this presentation, we first review classical approaches for validating mean-calibration, and we discuss the Likelihood Ratio Test (LRT) within the Exponential Dispersion Family (EDF). Then, we investigate the framework of universal inference to test for mean-calibration. We develop a sub-sampled split LRT within the EDF that provides finite sample guarantees with universally valid critical values. We investigate type I error, power and e-power of this sub-sampled split LRT, and we compare it to the classical LRT. We propose a novel test statistics based on a sub-sampled split Lq-Likelihood Ratio Test (LqRT) to enhance the performance of the calibration test. A numerical analysis verifies that universal inference with the sub-sampled split LRT and the sub-sampled split LqRT is an attractive alternative to the classical LRT achieving a high power in detecting miscalibration in medium and large samples.

Keywords: Calibration of mean predictions, e-values, e-power, Exponential Dispersion Family, isotonic regression, split Likelihood Ratio Test, universal inference.

[*]E-mail address: l.delong@uw.edu.pl
[†]E-mail address: mario.wuethrich@math.ethz.ch