THE UNIVERSITY OF CHICAGO | DEPARTMENT OF STATISTICS

PhD Dissertation Proposal Presentation

Raphael Rossellini
Department of Statistics
The University of Chicago

"Testing and ensuring calibration for decision-makers"

December 15, 2025, at 11:00 AM
DSI 322, 5460 S. University Ave.

Abstract

Forecast probabilities often serve as critical inputs for binary decision making. In such settings, calibration—ensuring forecasted probabilities match empirical frequencies—is essential. Although the common notion of Expected Calibration Error (ECE) provides actionable insights for decision making, it is not testable: it cannot be empirically estimated in many practical cases. Conversely, the recently proposed Distance from Calibration (dCE) is testable, but it is not actionable since it lacks decision-theoretic guarantees needed for high-stakes applications. To resolve this question, we consider Cutoff Calibration Error, a calibration measure that bridges this gap by assessing calibration over intervals of forecasted probabilities. We show that Cutoff Calibration Error is both testable and actionable, and we examine its implications for popular post-hoc calibration methods, such as isotonic regression and Platt scaling. We also identify common cases for which we can guarantee that isotonic regression yields a small Expected Calibration Error, which redounds to yet stronger decision-theoretic guarantees. Finally, we consider how to evaluate probabilistic models beyond calibration. A popular method is the energy score, which has at times been criticized for not being appropriately sensitive to incorrectly modeling correlation structures and marginal variances. We demonstrate that there exist simple modifications to the energy score that increase its sensitivity to these factors while retaining its original desirable characteristics.
