Colorful Pinball: Density-Weighted Quantile Regression for Conditional Guarantee of Conformal Prediction

Qianyi Chen, Bo Li* (Tsinghua University)

Scan for paper:

Conformal Prediction

Conformal prediction converts any predictive model into a set-valued predictor with distribution-free finite-sample coverage. Using calibration data, it selects a threshold on nonconformity scores and includes all candidate labels whose scores are sufficiently small:

$$\mathcal{C}_\alpha(x) = \{y: \ s(x,y) \le \widehat{q}_{1-\alpha}\},$$

$$1-\alpha \le \mathbb{P}\{Y \in \mathcal{C}_\alpha(X)\} \le 1-\alpha+\frac{1}{n-1}.$$

The nonconformity score $s(x,y)$ measures how incompatible a candidate label y is with the model prediction at x; for example, in regression, one may use the classical residual score $s(x,y) = |y - f(x)|$.

From Marginal to Conditional Coverage

We seek coverage that holds for each specific input X, known as conditional coverage. Our target is:

$$\mathbb{P}(Y \in \mathcal{C}_\alpha(X) \mid X) \text{ concentrate at } 1-\alpha.$$

We formulate it as the MSE of coverage probabilities, which is, however, difficult to evaluate and optimize.

$$\text{MSCE} := (\mathbb{P}(Y \in \mathcal{C}_\alpha(X) \mid X) - (1-\alpha))^2$$

Quantile Rectified Nonconformity Scores

To improve conditional coverage, we need a refined score component that captures local heteroscedastic structure. We therefore rectify the score by its estimated conditional $1-\alpha$ quantile, aiming to make coverage more uniform over X. We denote r as our final (rectified) score.

$$r(x,y) = s(x,y) - \hat{q}_{1-\alpha,s}(x)$$

Recover Hidden Heteroscedastic Structure

Standard practice estimates the quantile function of the score s via quantile regression, i.e., ERM with the pinball loss. Our derivation shows that, when targeting the MSE of conditional coverage probabilities, the plain pinball loss should be density-weighted.

$$\text{MSCE} \approx 2\mathbb{E}_X[f_{S|X}(q_{1-\alpha}(X) \mid X)\mathcal{E}(X)]$$

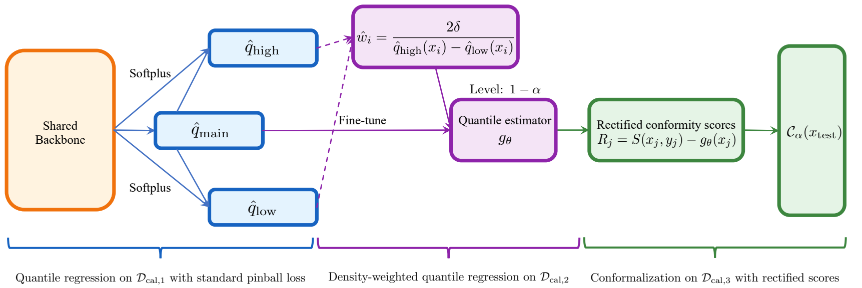
Density weighting amplifies heteroscedasticity and enables estimation by reciprocal finite differences instead of density estimation.

The excess pinball risk is the standard quantile-regression objective at X. Density reweighting reveals the full heteroscedastic spectrum.

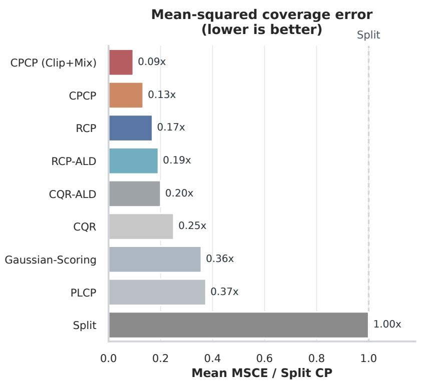
The key structure for bypassing density estimation: the quantile function is the inverse function of the CDF.

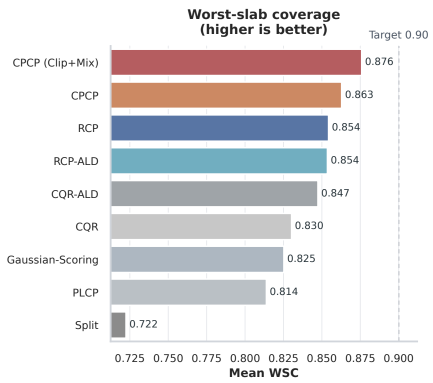
$$\frac{\partial q_\tau(x)}{\partial \tau} = \frac{1}{f_{S|X}(q_\tau(x))} \qquad \hat{w}(x) = \frac{2\delta}{\hat{q}_{\tau+\delta}(x) - \hat{q}_{\tau-\delta}(x)}$$

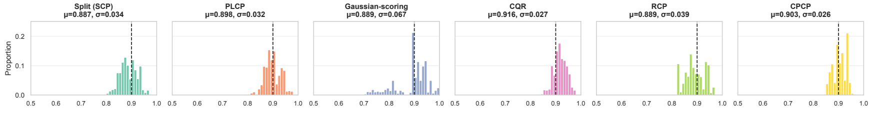
Flow of Our Methodology



Experimental Results







Estimated conditional coverage probabilities (by k-neighbors)