Conformal Inference under High-Dimensional Covariate Shifts via Likelihood-Ratio Regularization

Sunay Joshi*, Shayan Kiyani*, George Pappas,
Edgar Dobriban, Hamed Hassani

Penn
University of Pennsylvania

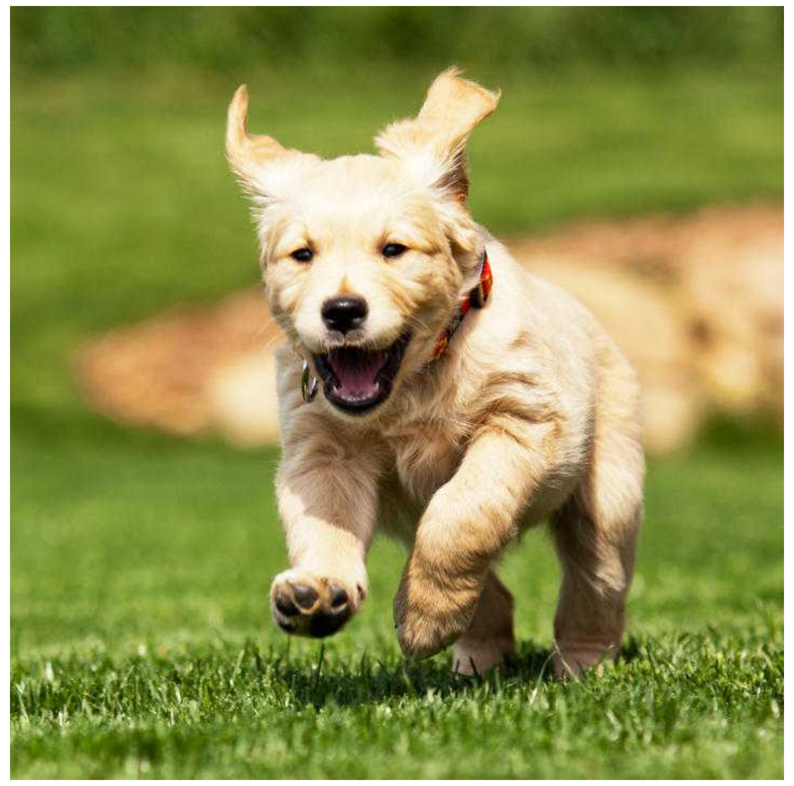
Setting: Conformal prediction for uncertainty quantification.

Goal: Given calibration samples (X_i, Y_i), for X_{test}, construct prediction set $\hat{C}(X_{test})$ with

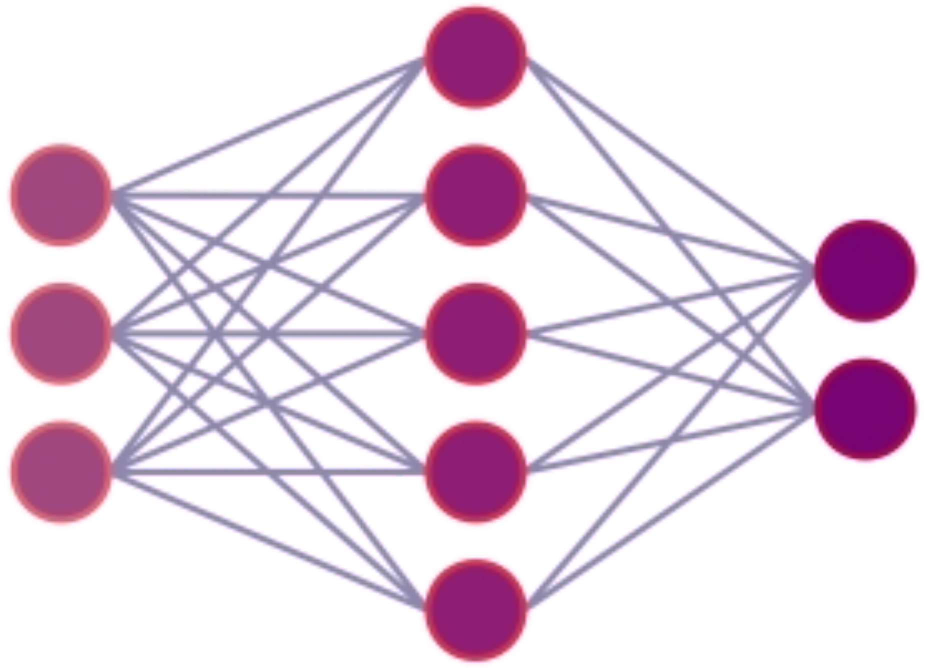
$$\mathbb{P}(Y_{test} \in \hat{C}(X_{test})) \geq 1 - \alpha \qquad \text{(marginal coverage)}$$

Learn threshold function $q(x)$ for score: $\hat{C}(x) = \{y : S(x, y) \leq q(x)\}$

Assumption: **exchangeability** of calibration + test samples



Input x



$\{$ dog, cat, tree $\}$

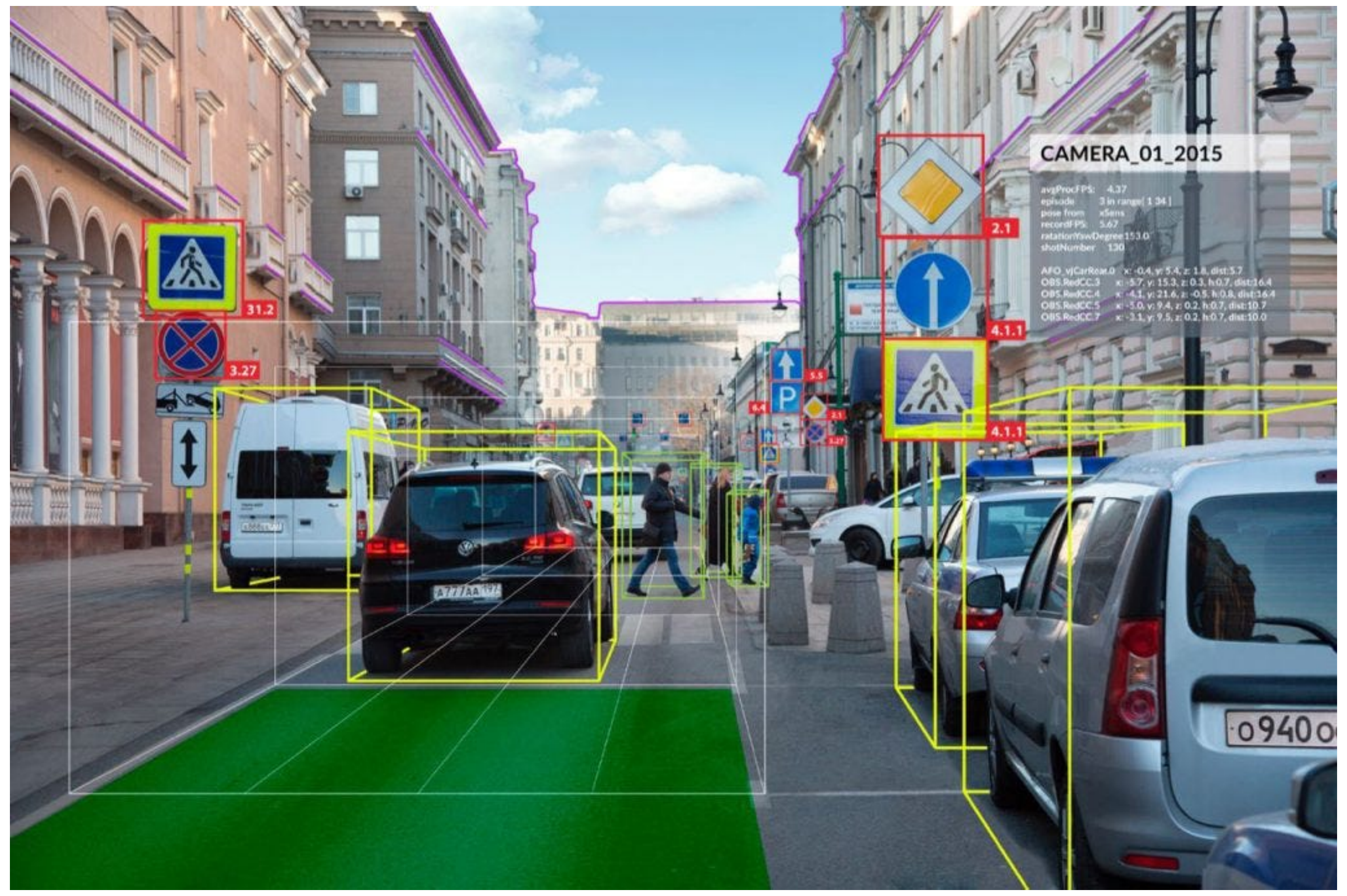
y^*

Prediction set $\hat{C}(x)$

Problem: How to address covariate shift?

Marginal distribution of features changes: $\mathbb{P}_{cal,X} \neq \mathbb{P}_{test,X}$

Conformal coverage guarantee may no longer hold



calibration

?

test

E.g., daytime vs nighttime images

Prior work: Quantile regression.

QR: minimize pinball loss over linear hypothesis class (Jung et al. '23, Gibbs et al. '25)

$$h^* = \arg\min_{h \in \mathcal{H}} \mathbb{E}_{cal}[\ell_\alpha(h(X), S(X, Y))]$$

Result: h^* has valid marginal test coverage for **any** likelihood-ratio $g \in \mathcal{H}$:

$$C(x) = \{y : S(x, y) \leq h^*(x)\} \implies \mathbb{Q}(S \leq h^*(X)) \geq 1 - \alpha,$$

$$\mathbb{Q} = \mathbb{Q}_X \times \mathbb{P}_{cal,Y|X}$$

$$d\mathbb{Q}_X / d\mathbb{P}_{cal,X} = g$$

(Justification: first-order conditions of loss w.r.t. perturbation g)

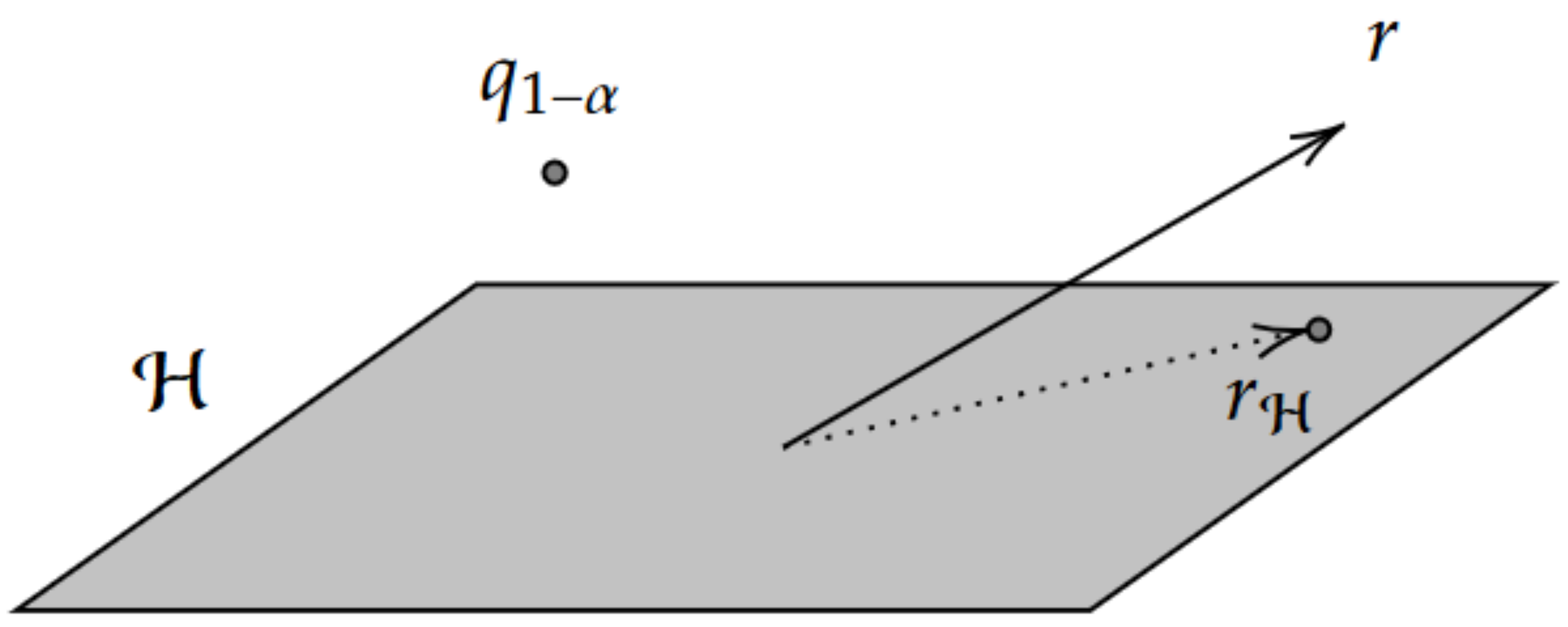
$$\frac{\partial}{\partial \epsilon}\Big|_{\epsilon=0} \mathbb{E}[\ell_\alpha(h^* + \epsilon g, S)] = \mathbb{E}_1[g(\mathbf{1}[S \leq h^*] - (1 - \alpha))]$$

Motivation for likelihood-ratio regularization.

Two ways to attain test coverage via QR:

$q_{1-\alpha} \in \mathscr{H}$ (conditional coverage)

$r \in \mathscr{H}$ (via first-order conditions)



Can we modify the objective to adapt to the two scenarios?

Our idea: likelihood-ratio regularization.

LR-QR (**L**ikelihood-Ratio **R**egularized **Q**uantile **R**egression): compute h^* via

$$(h^*, \beta^*) = \arg \min_{h \in \mathcal{H}, \beta \in \mathbb{R}} \mathbb{E}_{cal}[\ell_\alpha(h(X), S(X, Y))] + \lambda(\mathbb{E}_{cal}[\beta^2 h(X)^2] - \mathbb{E}_{test}[2\beta h(X)])$$

↑

via $\mathbb{E}_{cal}[(\beta h(X) - r(X))^2]$

(In finite samples, **no need to estimate** r; sample averages over labeled calibration data + **unlabeled** test data)

Extreme cases:

$\lambda = 0$ ———————————— $\lambda = \infty$

$\mathbb{E}_{cal}[\ell_\alpha(h(X), S(X, Y))]$

$\rightarrow h \approx q_{1-\alpha}$

$\mathbb{E}_{cal}[(\beta h(X) - r(X))^2]$

$\rightarrow h \propto r_{\mathcal{H}}$

Theory: marginal test coverage of LR-QR.

Writing $\hat{C}(x) = \{y : S(x, y) \leq \hat{h}(x)\}$ for LR-QR prediction set,

Theorem: given $\delta > 0$, for sufficiently many samples, w.p. $\geq 1 - \delta$,

$$\mathbb{P}_{test}(Y \in \hat{C}(X)) \geq (1 - \alpha) + 2\hat{\beta}\lambda\mathbb{E}_{cal}[(r_{\mathscr{H}}(X) - \hat{\beta}\hat{h}(X))^2] - \mathscr{E}_{cov} - (1 - \alpha)\mathbb{E}_{cal}[\,|\,r(X) - r_{\mathscr{H}}(X)\,|\,]$$

where $\mathscr{E}_{cov} = o(1)$.

Analyze LR-QR's data-dependent regularization via conditioning argument + stability

Experiments: Communities and Crime.

Predict violent crime rate from 127-dimensional socioeconomic feature vector

Compared to baselines, LR-QR more closely tracks nominal coverage

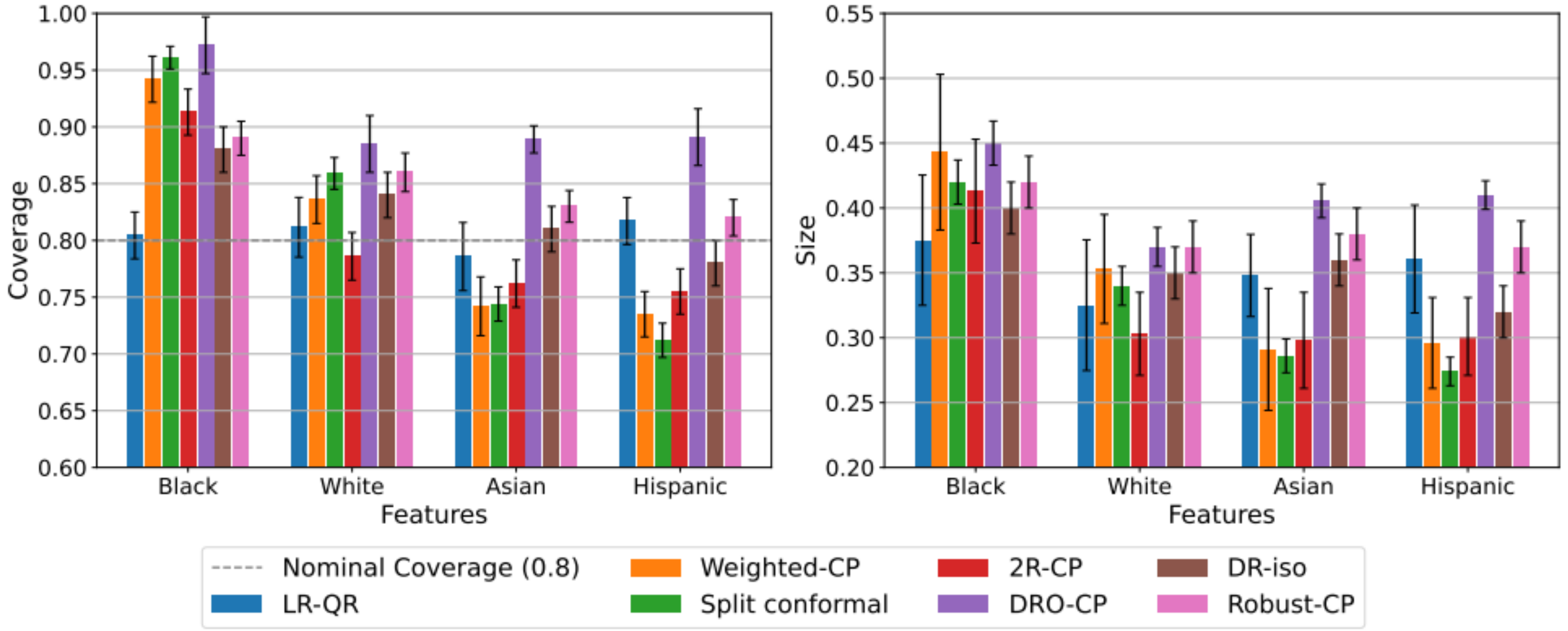


Figure 1: (Left) Coverage. (Right) Average prediction set size on the Communities and Crime dataset.