Testing for outliers with conformal p-values

Bates, Stephen, et al. The Annals of Statistics (2023)

Chanmoo Park

Seoul National University

February 25, 2025

Outline

Introduction

Goal: Outlier Detection with the **"Type-I error Guarantee"**

- In-distribution data : $\mathcal{D} = \{X_1, \ldots, X_{2n}\} \overset{\text{i.i.d.}}{\sim} P_X$.
- "Split" Conformal Prediction
  - $\mathcal{D}_{\text{train}} = \{X_1, \ldots, X_n\}$
  - $\mathcal{D}_{\text{cal}} = \{X_{n+1}, \ldots, X_{2n}\}$
- Test points: X_{new} or $\mathcal{D}_{\text{test}} = \{X_{2n+1}, \ldots, X_{2n+m}\}$ (Unknown distribution)
- Generally, in Split Confromal literature, $\mathcal{D}_{\text{train}}$ is considered as fixed (after training the baseline model, and we freeze the model).
- So we focus on the randomness on $\mathcal{D}_{\text{cal}}$, X_{new} and $\mathcal{D}_{\text{test}}$.

Marginally Conformal p-values (Naive version)

- Score function $\hat{s} : \mathbb{R}^d \to \mathbb{R}$ as a raw output of One-class classifier.
  Ex) One-Class SVM : a continuous output that small value imply outlier.
- For calibration set $\mathcal{D}^{\text{cal}}$, compute $\hat{s}(X_i)$ for $i = n+1, \dots, 2n$.
  Ex) Just evaluate OC-SVM on X_is.
- For a new test point X_{new}, define **"Marginally Conformal p-value"**

$$\hat{u}^{(\text{marg})}(X_{\text{new}}) = \frac{1 + \left|\{\, i : \hat{s}(X_i) \le \hat{s}(X_{\text{new}})\}\right|}{n+1}.$$

- Intuition 1:
  - $\hat{u}^{(\text{marg})}(X_{\text{new}})$ is based on the rank of the new score among the calibration scores.
  - $\hat{u}^{(\text{marg})}(X_{\text{new}})$ is uniformly distributed on $\{\frac{1}{n+1}, \frac{2}{n+1}, \dots, 1\}$.
    (under null hypothesis $H_0 : X_{\text{new}} \sim P_X$)

Marginally Conformal p-values (Naive version)

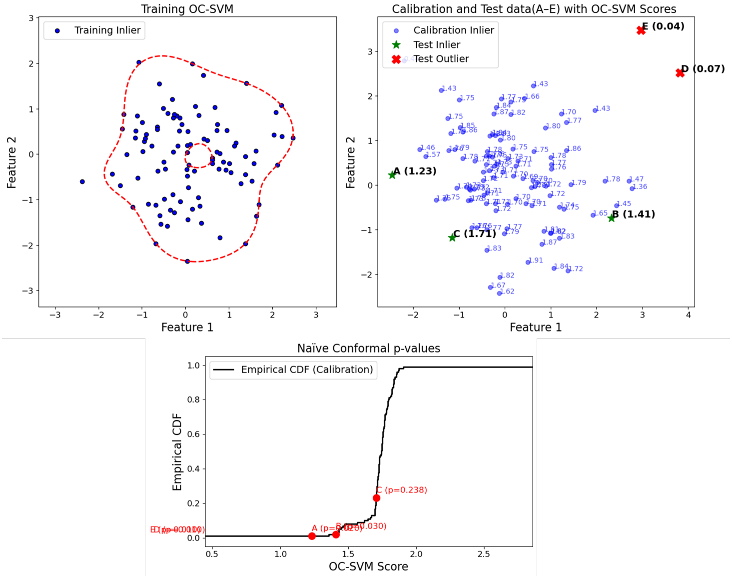
$$\hat{u}^{(\mathrm{marg})}(X_{\mathsf{new}}) \;=\; \frac{1 + \left|\{\, i : \hat{s}(X_i) \;\le\; \hat{s}(X_{\mathsf{new}})\}\right|}{n+1}.$$

- Marginal guarantee of $\hat{u}^{(\mathrm{marg})}$

$$\mathbb{P}_{\mathcal{D}_{\mathsf{cal}}, X_{\mathsf{new}}}\left[\hat{u}^{(\mathrm{marg})}(X_{\mathsf{new}}) \le \alpha\right] \;\le\; \alpha.$$

  - Test statistics : $\hat{s}(X_{\mathsf{new}})$
  - p-value under H_0 : $\hat{u}^{(\mathrm{marg})}(X_{\mathsf{new}})$
  - Reject H_0 when p-value is under α
  - (Marginal) Type I error : $\mathbb{P}_{\mathcal{D}_{\mathsf{cal}}, X_{\mathsf{new}}}\left[\hat{u}^{(\mathrm{marg})}(X_{\mathsf{new}}) \le \alpha\right]$
- Only guranteed "on average" over all possible $\mathcal{D}_{\mathsf{cal}}$.
- If you have a "unlucky" calibration set, it might not be guaranteed.

Marginally Conformal p-value (2-dimensional Toy example)



Calibration conditional p-values

Goal: Outlier Detection with the **"Conditional Type-I error Guarantee"**

- Marginal p-value & rejection region

$$\hat{u}^{(\mathrm{marg})}(X_{\mathsf{new}}) = \frac{1 + \left|\{ i : \hat{s}(X_i) \leq \hat{s}(X_{\mathsf{new}})\}\right|}{n+1}.$$

$$\hat{u}^{(\mathrm{marg})}(X_{\mathsf{new}}) \leq \alpha$$

- Problem: Our rejection procedure is highly dependent on $\mathcal{D}_{\mathsf{cal}}$
- Key Idea: Construct a universal envelope function $h(\cdot)$ for $\hat{u}^{(\mathsf{marg})}$.

Recall: $\hat{u}^{(\mathrm{marg})}(X_{\mathsf{new}})$ is uniformly distributed on $\{\frac{1}{n+1}, \frac{2}{n+1}, \dots, 1\}$.
(under null hypothesis $H_0 : X_{\mathsf{new}} \sim P_X$)

Calibration conditional p-values

- Since $\hat{u}^{(\text{marg})}$ follows uniform distribution, there is a few known ways to make an envelop function $h(\cdot)$ for uniform distribution.

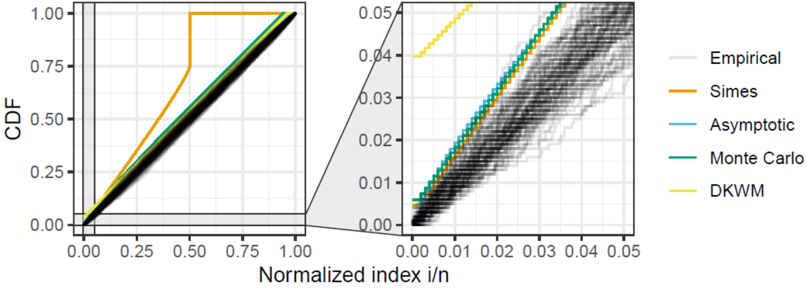


Figure: Envelope function for uniform CDF (Bates et al., 2023)

Calibration conditional p-values

- With $h(\cdot)$, define

$$\hat{u}^{(\mathrm{ccv})}(X_{\mathsf{new}}) \; = \; h\Big(\hat{u}^{(\mathrm{marg})}(X_{\mathsf{new}})\Big).$$

- **Calibration Conditional Validity** (guarantee)

$$\mathbb{P}_{\mathcal{D}_{\mathsf{cal}}}\left[\mathbb{P}_{X_{\mathsf{new}}}\left[\hat{u}^{(\mathrm{ccv})}\left(X_{\mathsf{new}}\right) \leq t \mid \mathcal{D}_{\mathsf{cal}}\right] \leq \alpha\right] \geq 1 - \delta$$

  - $\hat{u}^{(\mathsf{ccv})}$ is guaranteed at least for with at least $1 - \delta$ probability over the choice of $\mathcal{D}_{\mathsf{cal}}$.
  - That is, your unlucky choice of $\mathcal{D}_{\mathsf{cal}}$ is controlled under the probability of δ.

- For multiple testing $\mathcal{D}_{\mathsf{test}}$, BH (Benjamin-Hochberg) procedure is directly applicable.

Calibration conditional p-values (example)

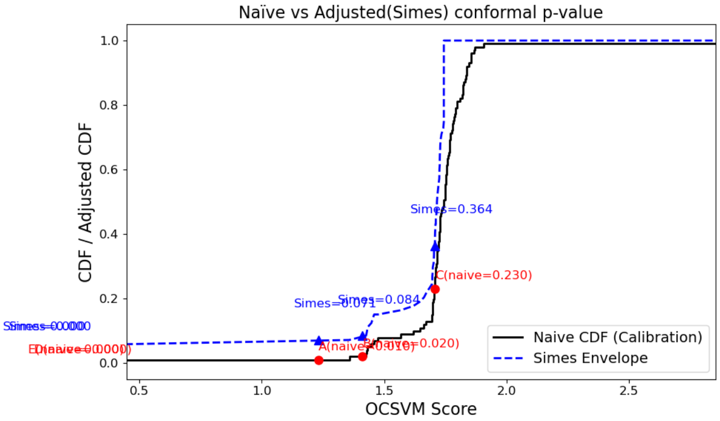


Figure: 2-dimensional toy example: Simes adjustment inflates the p-values.

Calibration conditional p-values (experiments)

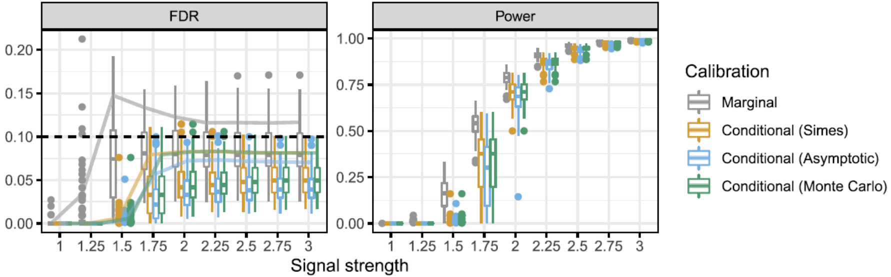


Figure: Conditinal FDR Simulation (Bates et al., 2023)

- $n_{\text{train}} = n_{\text{cal}} = n_{\text{test}} = 1000$,
- Conditional FDR: For a given calibration set, assess FDR of 100 different test sets
- Solid line 90th quantile of the conditional FDR. (100 experiments)