Conditional Testing based on Localized Conformal p-values

Presenter: Lin Lu

School of Statistics and Data Sciences, Nankai University, Tianjin, China

Email: linlu102099@gmail.com

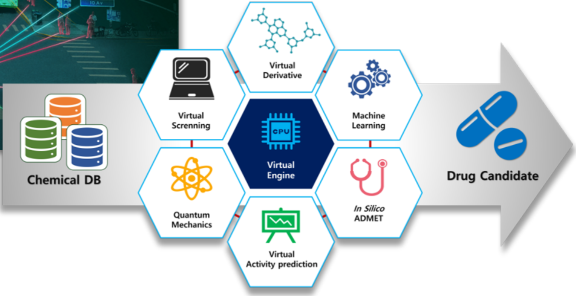
Joint work with
Xiaoyang Wu, Zhaojun Wang and Changliang Zou

ICLR 2025

March 10, 2025

Outline

ML predictions assist both daily and scientific tasks.



Problem set-up of conformal inference

Problem set-up

- Access to a labeled dataset $\mathcal{D} = \{(X_i, Y_i)\}_{i=1}^n \subset \mathcal{X} \times \mathcal{Y}$
- A test data point X_{n+1} with its label/response Y_{n+1} unobserved
- $\{(X_i, Y_i)\}_{i=1}^{n+1}$ are i.i.d.

- Train a prediction model $\widehat{\mu}$ on $\mathcal{D}$, then predict $\widehat{Y}_{n+1} = \widehat{\mu}(X_{n+1})$
- How to quantify the prediction uncertainty $\widehat{Y} \leftrightarrow Y$?

$\Longrightarrow$ To construct a **prediction interval** $\mathrm{PI}(\cdot)$ for X_{n+1} such that

$$\Pr(Y_{n+1} \in \mathrm{PI}(X_{n+1})) \geq 1 - \alpha.$$

Split conformal

The simplest method: **Split/Inductive Conformal** (Papadopoulos et al., 2002):

- Randomly split $\mathcal{D}$ into two equal-sized disjoint sets: training set $\mathcal{D}_{\mathcal{T}}$ and calibration set $\mathcal{D}_{\mathcal{C}}$ with index sets $\mathcal{T}$ and $\mathcal{C}$.
- Train predictive model $\hat{\mu}$ on training set $\mathcal{D}_{\mathcal{T}}$
- Compute the non-conformity scores $V_i = V(X_i, Y_i)$ (e.g., $V(x, y) = |y - \widehat{\mu}(x)|$) on calibration set $\mathcal{D}_{\mathcal{C}}$

- Construct prediction interval by thresholding:

$$\mathrm{PI}(X_{n+1}) = \{y : V(X_{n+1}, y) \leq \tau\}.$$

Split conformal

- To achieve finite sample coverage, take τ as the adjusted sample quantile

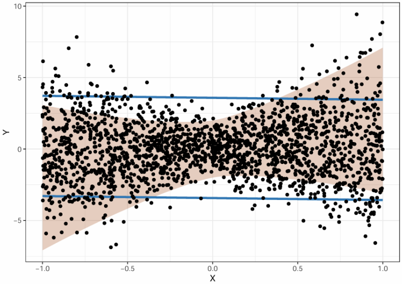
$$\widehat{q}_{\alpha} = Q\left(1-\alpha, \frac{1}{|\mathcal{C}|+1}\sum_{i\in\mathcal{C}}\delta_{V_i} + \frac{1}{|\mathcal{C}|+1}\delta_{\infty}\right).$$

$$\mathrm{PI}(X_{n+1}) = \{y : V(X_{n+1}, y) \le \widehat{q}_{\alpha}\}$$

- The coverage property holds since $\{V_i\}_{i\in\mathcal{C}} \cup \{V(X_{n+1}, Y_{n+1})\}$ are i.i.d.

$$\Pr(Y_{n+1} \in \mathrm{PI}(X_{n+1})) \ge 1-\alpha.$$

From marginal to conditional



- As a more appealing property, there could also be conditional valid prediction interval:

$$\Pr(Y_{n+1} \in \mathrm{PI}(X_{n+1}) \mid X_{n+1}) \geq 1 - \alpha.$$

However, this is impossible in the distribution-free context. (Lei and Wasserman, 2013)

- Recent works proposed different methods to improve conditional/local coverage. (Guan, 2023; Gibbs et al., 2023)

Localized conformal prediction

- Guan (2023):

$$\widehat{q}^{*}_{\widehat{\alpha},\mathrm{L}} = Q\left(1-\widehat{\alpha}; \sum_{i\in\mathcal{C}} H^{*}(X_i, X_{n+1})\delta_{V_i} + H^{*}(X_{n+1}, X_{n+1})\delta_{\infty}\right),$$

where $H^{*}(x,x') = \frac{H(x,x')}{\sum_{k\in\mathcal{C}} H(X_k,X_{n+1})+H(X_{n+1},X_{n+1})}$ for some kernel function $H(\cdot,\cdot)$ characterizing the similarity between its two arguments.

- Hore and Barber (2023):

First samples $\tilde{X}_{n+1}$ from the distribution $H(X_{n+1},\cdot)$, takes the threshold as

$$\widehat{q}_{\alpha,\mathrm{L}} = Q\left(1-\alpha; \sum_{i\in\mathcal{C}} \tilde{H}^{*}(X_i, \tilde{X}_{n+1})\delta_{V_i} + \tilde{H}^{*}(X_{n+1}, \tilde{X}_{n+1})\delta_{\infty}\right),$$

This again ensures finite sample coverage:

$$\Pr(Y_{n+1} \in \mathrm{PI}(X_{n+1})) = \Pr(V(X_{n+1}, Y_{n+1}) \leq \widehat{q}_{\alpha,\mathrm{L}}) \geq 1-\alpha.$$

From prediction interval to testing

- Confidence interval $\Longrightarrow$ Hypothesis testing
- Conformal prediction interval $\Longrightarrow$ Conformal p-value
- A valid prediction interval yields a valid conformal p-value by duality

$$p = \frac{\sum_{i \in \mathcal{C}} \mathbb{I}\{V_{n+1} \le V_i\} + 1}{|\mathcal{C}| + 1}.$$

  super-uniform when $\{(X_i, Y_i)\}_{i \in \mathcal{C}} \cup \{(X_{n+1}, Y_{n+1})\}$ are i.i.d.

- Applications on single hypothesis testing and multiple testing.

Conformal outlier detection

Bates et al. (2023) utilized conformal p-values to test for outliers.

- Access to clean data $\mathcal{D}_1 = \{X_{1i}\}_{i=1}^n \sim P_X$ and test data $\mathcal{D}_2 = \{X_{2j}\}_{j=1}^m$ with potential outliers

$$\mathbb{H}_{0j} : X_{2j} \sim P_X, \quad \text{versus} \quad \mathbb{H}_{1j} : \ \text{otherwise (outlier)},$$

- Split $\mathcal{D}_1$ in to $\mathcal{D}_1 = \mathcal{D}_{\mathcal{T}} \cup \mathcal{D}_{\mathcal{C}}$ with index sets $\mathcal{T}$ and $\mathcal{C}$
- Train a one-class classifier on $\mathcal{D}_{\mathcal{T}}$ as the score function $V(\cdot)$
- Compute scores on $\mathcal{D}_{\mathcal{C}}$ and $\mathcal{D}_2$, construct CP for $X_{2j} \in \mathcal{D}_2$

$$p_j = \frac{\sum_{i \in \mathcal{C}} \mathbb{I}\{V_{2j} \le V_{1i}\} + 1}{|\mathcal{C}| + 1}$$

- $(p_j)_{j=1}^m$ are PRDS, so applying BH procedure leads to finite sample FDR control.

Conformal testing

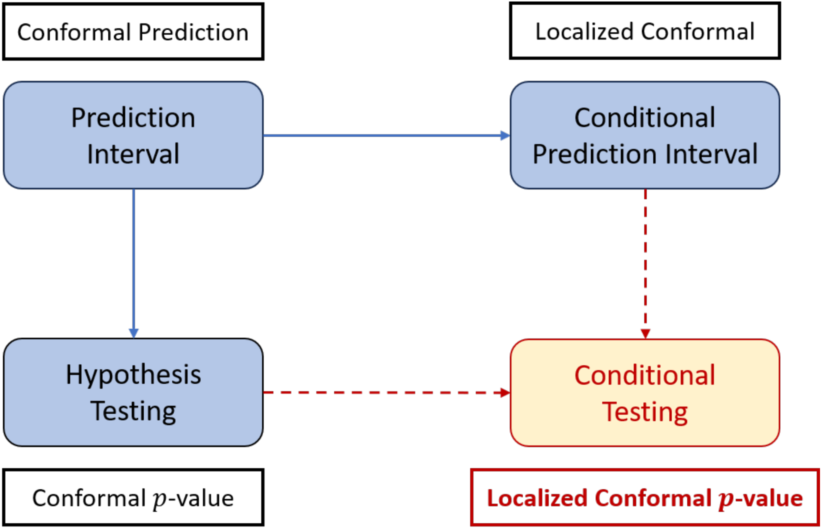
Other works

- Outlier detection: Zhang et al. (2022), Marandon et al. (2024), Liang et al. (2024)
- Data selection/sampling: Jin and Candès (2023), Wu et al. (2023)

Motivation

- Motivated by the capability of localized conformal prediction to capture local (conditional) information, we invert these prediction intervals to construct the localized conformal p-value (LCP).
- We consider several applications of the LCP, encompassing various conditional testing problems with different error criteria (e.g., FDR, FWER or type I error).

Our work


Conformal Prediction
Localized Conformal
Prediction Interval
Conditional Prediction Interval
Hypothesis Testing
Conditional Testing
Conformal p-value
Localized Conformal p-value

Outline for LCP: Localized conformal p-values

Localized conformal p-values

- $\mathcal{D}_1 = \{(X_{1i}, Y_{1i})\}_{i=1}^{n} \overset{iid}{\sim} P_{1,X} \times P_1 = P_{1,Y|X}$, split into $\mathcal{D}_1 = \mathcal{D}_{\mathcal{T}} \cup \mathcal{D}_{\mathcal{C}}$
- $\mathcal{D}_2 = \{(X_{2j}, Y_{2j})\}_{j=1}^{m}$ with each $(X_{2j}, Y_{2j}) \sim P_{2j} = P_{2,X} \times P_{2j,Y|X}$.
- Score function V obtained on $\mathcal{D}_{\mathcal{T}}$

Definition: Localized conformal p-value (LCP)

Let $H(x, x') = \frac{1}{h^d} K\left(\frac{x-x'}{h}\right)$ be a kernel function. The localized conformal p-value for $(X_{2j}, Y_{2j}) \in \mathcal{D}_2$ is defined as

$$p_{\mathrm{L},j} = \frac{\sum_{i\in\mathcal{C}} H(X_{1i}, \tilde{X}_{2j})\mathbb{I}\{V_{2j} \le V_{1i}\} + \xi_j \cdot H(X_{2j}, \tilde{X}_{2j})}{\sum_{i\in\mathcal{C}} H(X_{1i}, \tilde{X}_{2j}) + H(X_{2j}, \tilde{X}_{2j})},$$

where $\xi_j \sim U[0,1]$ is an independent random variable and $\tilde{X}_{2j} \sim H(X_{2j}, \cdot)$.

- A localized counterpart of the conformal p-value

Basic properties of LCP

Theorem (Finite-sample validity)

Under the condition $P_1 = P_{2j}$, the localized conformal p-value satisfies

$$\Pr(p_{\mathrm{L},j} \le \alpha) \le \alpha, \quad 0 \le \alpha \le 1.$$

Furthermore, if the score V has a continuous distribution, then

$$\Pr(p_{\mathrm{L},j} \le \alpha) = \alpha, \quad 0 \le \alpha \le 1.$$

Theorem (Covariate shift)

If $P_{1,X} \ne P_{2,X}$, denote the covariate density ratio as $g(\mathbf{x}) := \frac{\mathrm{d}P_{2,X}}{\mathrm{d}P_{1,X}}(x)$ then

$$\Pr(p_{\mathrm{L},j} \le \alpha) \le \alpha + \|f_{1,X}\|_\infty \mathbb{E}_{X \sim P_{H,X}, U \sim K(\cdot)} \{|g(X + hU) - g(X)|\},$$

where the distribution $P_{H,X}$ has a density function $f_{H,X}(x) = \mathbb{E}_{X \sim P_{1,X}}\{H(X, x)\}$.

Basic properties of LCP

Assumption (1)

The following conditions hold for $(X, Y) \sim P_1$*:*

- $V(X, Y)$ *has a continuous distribution with bounded density;*
- *The conditional distribution of the score* $V = V(X, Y)$ *satisfies*

$$\|F_{V|X=x}(v) - F_{V|X=x'}(v)\|_\infty \leq L \cdot \|x - x'\|_2^\beta$$

  for some constant $L > 0, 0 < \beta \leq 1$*. That is, the conditional distribution function* $F_{V|X=x}$ *varies smoothly with* x*.*
- *The density function* $f_{1,X}(x)$ *is continuous, and the conditional density function* $f_1(y \mid x)$ *is continuous in* x*.*

Basic properties

Theorem

Define the LCP function as

$$p_{\mathrm{L}}(x,y) = \frac{\sum_{i\in\mathcal{C}} H(X_{1i},\tilde{X})\mathbb{I}\{v \le V_{1i}\} + \xi \cdot H(x,\tilde{X})}{\sum_{i\in\mathcal{C}} H(X_{1i},\tilde{X}) + H(x,\tilde{X})}.$$

Assume Assumption 1 holds, then the LCP function satisfies

$$|p_{\mathrm{L}}(x,y) - (1 - F_{V|X=x}(v))| = O_p\left(\sqrt{h^{2\beta} + \frac{1}{nh^d}}\right)$$

for any fixed (x,y) *with* $v = V(x,y)$*, as* $h \to 0, nh^d \to \infty$.

- By the weak law of large number, the unweighted CP function satisfies

$$p_{\mathrm{CP}}(x,y) = \frac{\sum_{i\in\mathcal{C}} \mathbb{I}\{v \le V_{1i}\} + 1}{|\mathcal{C}| + 1} \xrightarrow{p} 1 - F_V(v).$$

Outline for Applications on conditional testing

Conditional outlier detection: motivation

- For labeled data (X, Y), outliers in response Y is often more important.
- Outlyingness of Y depends on X, resulting in conditional outliers.

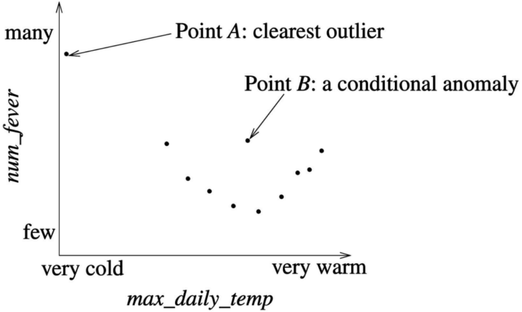


Figure: Illustration of conditional outliers in Song et al. (2007).

Conditional outlier detection: problem set-up

- $\mathcal{D}_1 = \{(X_{1i}, Y_{1i})\}_{i=1}^{n}$: clean data split into $\mathcal{D}_1 = \mathcal{D}_{\mathcal{T}} \cup \mathcal{D}_{\mathcal{C}}$
  $\mathcal{D}_2 = \{(X_{2j}, Y_{2j})\}_{j=1}^{m}$: test data with potential outliers.
- Detecting conditional outliers can be formulated as the multiple testing problem:

$$\mathbb{H}_{0j} : P_{2j,Y|X} = P_{1,Y|X}, \text{ versus } \mathbb{H}_{1j} : \text{ otherwise (outlier).}$$

- Score function V fitted on $\mathcal{D}_{\mathcal{T}}$:
  - residual: $V(x, y) = |y - \widehat{\mu}(x)|$
  - CQR score: $V(x, y) = \max\{\hat{q}_{lo}(x) - y, y - \hat{q}_{hi}(x)\}$

LCP for conditional outlier detection

- Compute scores $\{V_{1i}\}_{i\in\mathcal{C}}$ and $\{V_{2j}\}_{j=1}^{m}$ on $\mathcal{D}_{\mathcal{C}}$ and $\mathcal{D}_2$. Construct LCP's $\{p_{\mathrm{L},j}\}_{j=1}^{m}$ as

$$p_{\mathrm{L},j} = \frac{\sum_{i\in\mathcal{C}} H(X_{1i}, \tilde{X}_{2j})\mathbb{I}\{V_{2j} \leq V_{1i}\} + \xi_j \cdot H(X_{2j}, \tilde{X}_{2j})}{\sum_{i\in\mathcal{C}} H(X_{1i}, \tilde{X}_{2j}) + H(X_{2j}, \tilde{X}_{2j})}.$$

- However, the localized conformal p-values are no longer PRDS.
- Apply the conditional calibration technique (Fithian and Lei, 2022) to achieve finite sample FDR control.

Conditional calibration

- Compute auxiliary p-values $p_{\mathrm{L},j}^{(\ell)}$ (a simple modified version).

$$p_{\mathrm{L},j}^{(\ell)} = \frac{\sum_{i\in\mathcal{C}} H(X_{1i},\tilde{X}_{2j})\mathbb{I}\{V_{2j}\le V_{1i}\} + \xi_j\cdot H(X_{2j},\tilde{X}_{2j})\mathbb{I}\{V_{2j}\le V_{2l}\}}{\sum_{i\in\mathcal{C}} H(X_{1i},\tilde{X}_{2j}) + H(X_{2j},\tilde{X}_{2j})}.$$

- Let $\widehat{\mathcal{R}}_{j\to 0}$ be the rejection set of running BH procedure on $\{p_{\mathrm{L},1}^{(j)},\ldots,p_{\mathrm{L},j-1}^{(j)},0,p_{\mathrm{L},j+1}^{(j)},\ldots,p_{\mathrm{L},m}^{(j)}\}$ with nominal level α.
- Generating independent $\zeta_1,\ldots,\zeta_m\sim\mathrm{U}[0,1]$, determine the final rejection set by

$$\mathcal{R} = \left\{j : p_{\mathrm{L},j}\le \frac{\alpha|\widehat{\mathcal{R}}_{j\to 0}|}{m}, \zeta_j|\widehat{\mathcal{R}}_{j\to 0}|\le r^*\right\},$$

$$r^* = \max\left\{r : \sum_{j=1}^{m}\mathbb{I}\left\{p_{\mathrm{L},j}\le \alpha|\widehat{\mathcal{R}}_{j\to 0}|/m, \zeta_j|\widehat{\mathcal{R}}_{j\to 0}|\le r\right\}\ge r\right\}.$$

LCP for conditional outlier detection

Our detection procedure can guarantee finite-sample FDR control.

Theorem (Finite-sample FDR control)

Denote the inlier index set as $\mathcal{I} \subseteq [m]$. Under the condition that $P_{1,\mathbf{X}} = P_{2,\mathbf{X}}$, the final output $\mathcal{R}$ ensures

$$\mathrm{FDR} = \mathbb{E}\left(\frac{|\mathcal{R} \cap \mathcal{I}|}{|\mathcal{R}| \vee 1}\right) \leq \alpha.$$

Conditional label screening: problem set-up

- Consider a multi-response setting with $\mathbf{Y} = (Y_1, \ldots, Y_S)$ is a vector. The r.v. S is the length of $\mathbf{Y}$ (could be a constant).
- Aim to screen out components of $\mathbf{Y}$ not satisfying the pre-given rule $Y_s \in \mathcal{A}_s$ and keep the rest without observing $\mathbf{Y}$.
- LLM factuality: $Y_s = 1, 0 \Longrightarrow$ the claim is correct or not, $\mathcal{A}_s = \{1\}$.
- Medical diagnosis: Y_s's are health metrics, $\mathcal{A}_s = (-\infty, a]$ or $[b, +\infty)$

Conditional label screening

- Labeled data $\mathcal{D}_1 = \{(X_{1i}, S_{1i}, \mathbf{Y}_{1i})\}_{i=1}^{n}$
  Unlabeled data $\mathcal{D}_2 = \{(X_{2j}, S_{2j})\}_{j=1}^{m}$
- This can be formulated as a multiple testing problem for each test sample

$$\mathbb{H}_{0j,s}: \; Y_{2j,s} \notin \mathcal{A}_s, \; \text{ versus } \; \mathbb{H}_{1j,s}: \; Y_{2j,s} \in \mathcal{A}_s, \quad 1 \leq s \leq S_{2j}. \tag{1}$$

- Let $\delta_{j,s} = 1$ or 0 indicate reject $\mathbb{H}_{0j,s}$ or not, we seek to control the FWER of (1)

$$\text{FWER} = \Pr\left(\sum_{s=1}^{S_{2j}} \mathbb{I}\{Y_{2j,s} \notin \mathcal{A}_s, \delta_{j,s} = 1\} > 0\right) \leq \alpha,$$

LCP for conditional label screening

- We can still fit a score function $V(\cdot)$ on $\mathcal{D}_{\mathcal{T}}$ with larger value of $V_{2j,s} = V(X_{2j})_s$ indicating more evidence for $Y_{2j,s} \in \mathcal{A}_s$.
- Utilizing the conformal p-value leads to finite sample FWER control.
- However, the conditional error rate matters since the multiple testing problem is defined for each test point.

$$\text{cFWER} = \Pr\left(\sum_{s=1}^{S_{2j}} \mathbb{I}\{Y_{2j,s} \notin \mathcal{A}_s, \delta_{j,s} = 1\} > 0 \mid X_{2j}\right)$$

$\Longrightarrow$ use LCP to mitigate conditional FWER inflation

LCP for conditional label screening

- Define the LCP for each component $Y_{2j,s}$

$$p_{\mathrm{L},j,s} = \frac{\sum_{i\in\mathcal{C}} H(X_{1i},\tilde{X}_{2j})\mathbb{I}\{V_{2j,s} \le \bar{V}_{1i}\} + \xi_j \cdot H(X_{2j},\tilde{X}_{2j})}{\sum_{i\in\mathcal{C}} H(X_{1i},\tilde{X}_{2j}) + H(X_{2j},\tilde{X}_{2j})},$$

  where $\bar{V}_{1i} = \max\{V_{1i,s} : Y_{1i,s} \notin \mathcal{A}_s\}$.

- The $p_{\mathrm{L},j,s}$'s satisfy the group super-uniform property, i.e.,

$$\Pr\left(\bigcup_{Y_{2j,s}\notin\mathcal{A}_s} \{p_{\mathrm{L},j,s} \le \alpha\}\right) \le \alpha,$$

- Screening out components of $\mathbf{Y}_{2j}$ with $p_{\mathrm{L},j,s} \le \alpha$.

LCP for conditional label screening

Theorem

Suppose $\{(X_{1i}, S_{1i}, \mathbf{Y}_{1i})\}_{i\in\mathcal{C}} \cup \{(X_{2j}, S_{2j}, \mathbf{Y}_{2j})\}_{j=1}^{m}$ *are exchangeable, then the label screening procedure given by our procedure ensures finite-sample FWER control*

$$\mathrm{FWER} = \Pr\left(\sum_{s=1}^{S_{2j}} \mathbb{I}\{Y_{2j,s} \notin \mathcal{A}_s, p_{\mathrm{L},j,s} \le \alpha\} > 0\right) \le \alpha.$$

Moreover, for any fixed set $\mathcal{B} \subset \mathcal{X}$ *with* $\Pr(X_{2j} \in \mathcal{B}) > 0$, *the conditional FWER has the following bound*

$$\mathrm{cFWER}_{\mathcal{B}} = \Pr\left(\sum_{s=1}^{S_{2j}} \mathbb{I}\{Y_{2j,s} \notin \mathcal{A}_s, p_{\mathrm{L},j,s} \le \alpha\} > 0 \;\middle|\; X_{2j} \in \mathcal{B}\right)$$
$$\le \alpha + 2\|f_{1,X}\|_\infty \frac{\Pr_{X\sim P_{H,X}, U\sim K(\cdot)}(\|U\|_2 \ge h^{-1} d(X, \partial\mathcal{B}))}{\Pr(X_{2j} \in \mathcal{B})},$$

where $\partial\mathcal{B}$ *is the boundary of set* $\mathcal{B}$.

Outline for Numerical results

Set-up for conditional outlier detection

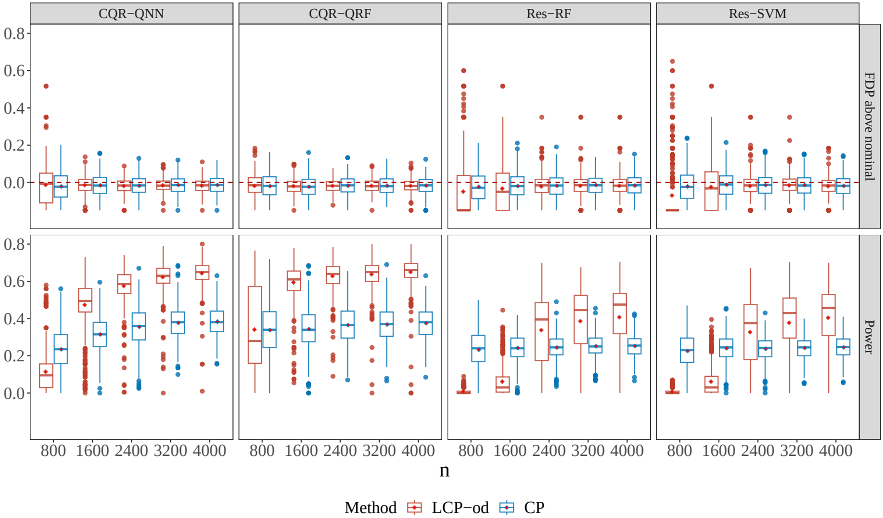
Scenario A1:

- $\mathbf{X} = (X_1, \ldots, X_{d-1})^\top \in \mathbb{R}^{d-1}$ with $d = 10$ and an additional time feature $t \in \mathbb{R}$.
- Inlier model:
$$Y = \mathbf{X}\boldsymbol{\beta} + (3 + 2 \cdot \sin(2\pi \cdot t)) \cdot \varepsilon$$
with $X_1, \ldots, X_{d-1} \sim \mathrm{U}[-1, 1]$, $t \sim \mathrm{U}[0, 1]$ and $\varepsilon \sim N(0, 1)$ independently.
- 10% outliers following the model
$$Y = \mathbf{X}\boldsymbol{\beta} + (3 + 2 \cdot \sin(2\pi \cdot t)) \cdot \varepsilon + r(t) \cdot \xi,$$
where $r(t) = 3 \cdot (3 + 1.5 \sin(2\pi \cdot t))$ and $\Pr(\xi = \pm 1) = 1/2$.
- The coefficient vector is $\boldsymbol{\beta} = (0.5, -0.5, 0.5, -0.5, 0.5, 0, 0, 0, 0)$.

Methods:

- Our method: LCP-od
- Bates et al. (2023): CP

Results for conditional outlier detection



Set-up for conditional label screening

Consider a nonlinear regression scenario (Scenario A2).

- Take a constant $S = 2$ and the response $\mathbf{Y} = (Y_1, Y_2)$ with
  $Y_1 = -2X_1 + 7X_2^2 + 3\exp(X_3 + 2X_4^2) + \varepsilon$,
  $Y_2 = -6X_1 + 5X_2^2 + 3\exp(2X_3 + X_4^2) + \varepsilon$, $\mathbf{X} \sim U[-1, 1]^4$ and $\varepsilon \sim \mathcal{N}(0, 1)$.
- The screening target is $Y_s \in \mathcal{A}_s = [a_s, +\infty)$ where a_s is the 70% quantile of Y_s for $s = 1, 2$, respectively.
- Fix the sample sizes $n = 500$, $m = 2000$ and vary $\alpha \in \{0.05, 0.1, 0.15, 0.2\}$.

Benchmarks.

- Compare our conditional label screening method via LCP (abbreviated as LCP-ls) with the thresholding procedure without weighting (abbreviated as THR).
- The THR method is performed by replacing LCP with the classical unweighted CP.

Results for conditional label screening

- (Marginal FWER): $\text{mFWER} = \Pr\left(\sum_{s=1}^{S_{2j}} \mathbb{I}\{Y_{2j,s} \notin \mathcal{A}_s, \delta_{j,s} = 1\} > 0\right)$
- (Conditional FWER): $\text{cFWER}_k = \Pr\left(\sum_{s=1}^{S_{2j}} \mathbb{I}\{Y_{2j,s} \notin \mathcal{A}_s, \delta_{j,s} = 1\} > 0 \mid \mathbf{X} \in \mathcal{B}_k\right)$ for 4 different regions $\mathcal{B}_1, \mathcal{B}_2, \mathcal{B}_3, \mathcal{B}_4$

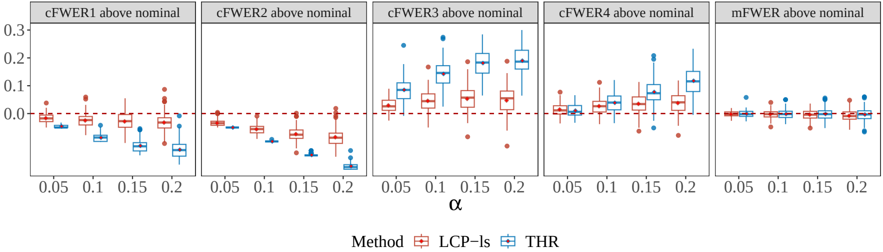


Figure: Conditional FWER (cFWER) and marginal FWER (mFWER) above nominal.

Summary

- We propose the localized conformal p-value by utilizing localization techniques in Guan (2023) and Hore and Barber (2023).
- The proposed LCP enjoys theoretical properties indicating its adequacy for conditional testing problems.
- We apply the LCP on several application problems with different preferred error criteria

Thank you!

See more applications, details and experiments results in our paper:

- Conditional Testing based on Localized Conformal p-values, *ICLR*, 2025.

Bates, S., Candès, E., Lei, L., Romano, Y., and Sesia, M. (2023), "Testing for outliers with conformal p-values," *The Annals of Statistics*, 51, 149–178.

Fithian, W. and Lei, L. (2022), "Conditional calibration for false discovery rate control under dependence," *The Annals of Statistics*, 50, 3091 – 3118.

Gibbs, I., Cherian, J. J., and Candès, E. J. (2023), "Conformal Prediction With Conditional Guarantees," .

Guan, L. (2023), "Localized conformal prediction: A generalized inference framework for conformal prediction," *Biometrika*, 110, 33–50.

Hore, R. and Barber, R. F. (2023), "Conformal prediction with local weights: randomization enables local guarantees," *arXiv preprint arXiv:2310.07850*.

Jin, Y. and Candès, E. J. (2023), "Selection by Prediction with Conformal p-values," *Journal of Machine Learning Research*, 24, 1–41.

Lei, J. and Wasserman, L. (2013), "Distribution-free Prediction Bands for Non-parametric Regression," *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 76, 71–96.

Liang, Z., Sesia, M., and Sun, W. (2024), "Integrative conformal p-values for out-of-distribution testing with labelled outliers," *Journal of the Royal Statistical Society Series B: Statistical Methodology*, qkad138.

Marandon, A., Lei, L., Mary, D., and Roquain, E. (2024), "Adaptive novelty detection with false discovery rate guarantee," *The Annals of Statistics*, 52, 157 – 183.

Papadopoulos, H., Proedrou, K., Vovk, V., and Gammerman, A. (2002), "Inductive confidence machines for regression," in *European Conference on Machine Learning*, New York: Springer, pp. 345–356.

Song, X., Wu, M., Jermaine, C., and Ranka, S. (2007), "Conditional Anomaly Detection," *IEEE Transactions on Knowledge and Data Engineering*, 19, 631–645.

Wu, X., Huo, Y., Ren, H., and Zou, C. (2023), "Optimal Subsampling via Predictive Inference," *Journal of the American Statistical Association*, 0, 1–29.

Zhang, Y., Jiang, H., Ren, H., Zou, C., and Dou, D. (2022), "AutoMS: Automatic Model Selection for Novelty Detection with Error Rate Control," in *Advances in Neural Information Processing Systems*, vol. 35, pp. 19917–19929.