Conformal Prediction with Missing Values

Margaux Zaffran

Stats Workshop (FAST-BIG)

Aymeric Dieuleveut
Ecole
Polytechnique
Paris - France

Julie Josse
INRIA
PreMeDICaL
Montpellier - France

Yaniv Romano
Technion - Israel Institute
of Technology
Haifa - Israel

(Way too short) Intro to (Split) Conformal Prediction

- Standard Split Conformal Prediction for Mean-Regression
- Improving Adaptiveness: Conformalized Quantile Regression

Conformal Prediction with Missing Values

Setting

- $(X, Y) \in \mathbb{R}^d \times \mathbb{R}$ random variables
- n training samples $\left(X^{(k)}, Y^{(k)}\right)_{k=1}^{n}$
- Goal: predict an unseen point $Y^{(n+1)}$ at $X^{(n+1)}$ with **confidence**
- How? Given a miscoverage level $\alpha \in [0, 1]$, build a predictive set $\mathcal{C}_\alpha$ such that:

$$\mathbb{P}\left\{Y^{(n+1)} \in \mathcal{C}_\alpha\left(X^{(n+1)}\right)\right\} \geq 1 - \alpha, \tag{1}$$

  and $\mathcal{C}_\alpha$ should be as small as possible, in order to be informative.
  For example: $\alpha = 0.1$ and obtain a 90% coverage interval

▶ Construction of the predictive intervals should be

  - agnostic to the model
  - agnostic to the data distribution
  - valid in finite samples

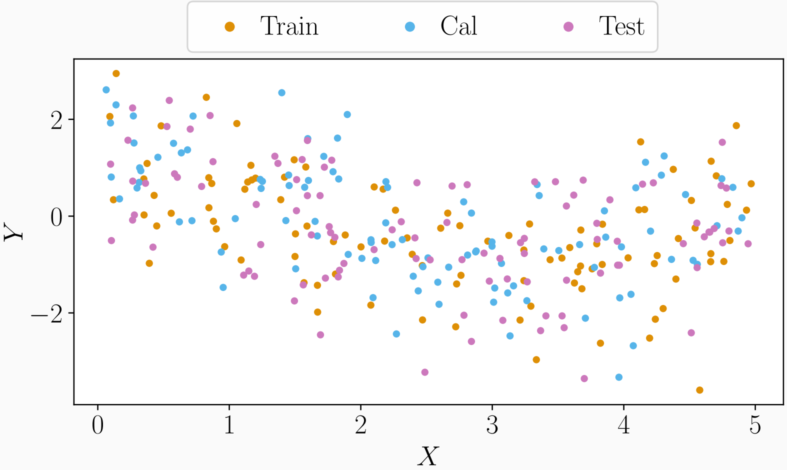
(Way too short) Intro to (Split) Conformal Prediction

Standard Split Conformal Prediction for Mean-Regression

Improving Adaptiveness: Conformalized Quantile Regression

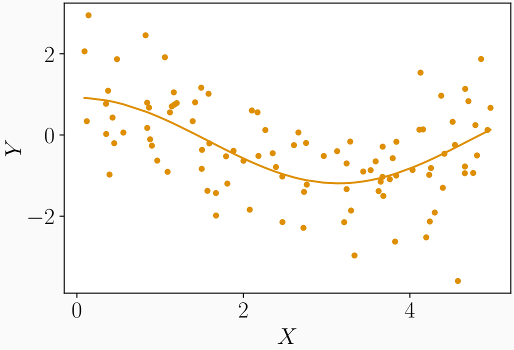
Conformal Prediction with Missing Values

Split Conformal Prediction (SCP)[1,2,3]: toy example



[1] Vovk et al. (2005), *Algorithmic Learning in a Random World*
[2] Papadopoulos et al. (2002), *Inductive Confidence Machines for Regression*, ECML
[3] Lei et al. (2018), *Distribution-Free Predictive Inference for Regression*, JRSS B

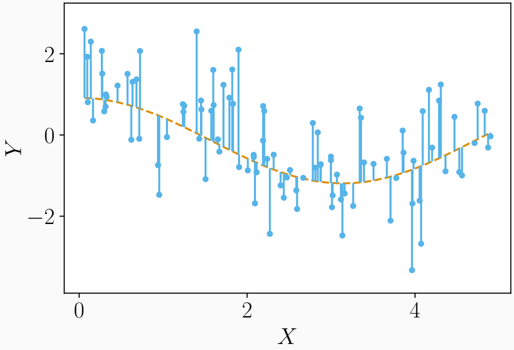
Split Conformal Prediction (SCP)[1,2,3]: training step



- Learn (or get) $\hat{\mu}$

[1] Vovk et al. (2005), *Algorithmic Learning in a Random World*
[2] Papadopoulos et al. (2002), *Inductive Confidence Machines for Regression*, ECML
[3] Lei et al. (2018), *Distribution-Free Predictive Inference for Regression*, JRSS B

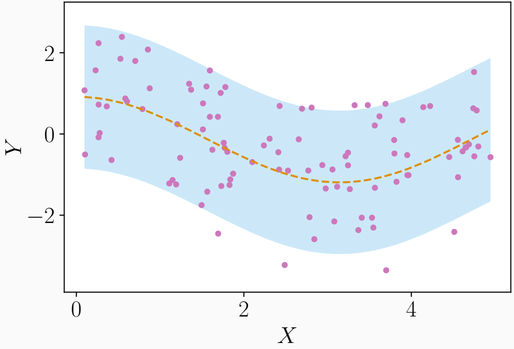
Split Conformal Prediction (SCP)[1,2,3]: calibration step



- Predict with $\hat{\mu}$
- Get the $|\text{residuals}|$, a.k.a. scores $\left\{S^{(k)}\right\}_{k\in\mathrm{Cal}}$
- Compute the $(1-\alpha)$ empirical quantile of $\mathcal{S} = \{|\text{residuals}|\}_{\mathrm{Cal}} \cup \{+\infty\}$, noted $q_{1-\alpha}(\mathcal{S})$

[1] Vovk et al. (2005), *Algorithmic Learning in a Random World*
[2] Papadopoulos et al. (2002), *Inductive Confidence Machines for Regression*, ECML
[3] Lei et al. (2018), *Distribution-Free Predictive Inference for Regression*, JRSS B

Split Conformal Prediction (SCP)[1,2,3]: prediction step



- Predict with $\hat{\mu}$
- Build $\widehat{C}_\alpha(x)$: $[\hat{\mu}(x) \pm q_{1-\alpha}(\mathcal{S})]$

[1] Vovk et al. (2005), *Algorithmic Learning in a Random World*
[2] Papadopoulos et al. (2002), *Inductive Confidence Machines for Regression*, ECML
[3] Lei et al. (2018), *Distribution-Free Predictive Inference for Regression*, JRSS B

SCP: theoretical foundation

Definition (Exchangeability)

$\left(X^{(k)}, Y^{(k)}\right)_{k=1}^{n}$ are exchangeable if for any permutation σ of $[\![1, n]\!]$ we have:

$$\begin{aligned} & \mathcal{L}\left(\left(X^{(1)}, Y^{(1)}\right), \ldots, \left(X^{(n)}, Y^{(n)}\right)\right) \\ = \quad & \mathcal{L}\left(\left(X^{(\sigma(1))}, Y^{(\sigma(1))}\right), \ldots, \left(X^{(\sigma(n))}, Y^{(\sigma(n))}\right)\right), \end{aligned}$$

where $\mathcal{L}$ designates the joint distribution.

Examples of exchangeable sequences

- i.i.d. samples
- The components of $\mathcal{N}\left(\begin{pmatrix} m \\ \vdots \\ \vdots \\ m \end{pmatrix}, \begin{pmatrix} \sigma^2 & & & \\ & \ddots & \gamma^2 & \\ & \gamma^2 & \ddots & \\ & & & \sigma^2 \end{pmatrix}\right)$

SCP: theoretical guarantees

SCP enjoys finite sample guarantees proved in Vovk et al. (2005); Lei et al. (2018).

Theorem

Suppose $\left(X^{(k)}, Y^{(k)}\right)_{k=1}^{n+1}$ *are exchangeable (or i.i.d.). SCP applied on* $\left(X^{(k)}, Y^{(k)}\right)_{k=1}^{n}$ *outputs* $\widehat{C}_\alpha(\cdot)$ *such that:*

$$\mathbb{P}\left\{Y^{(n+1)} \in \widehat{C}_\alpha\left(X^{(n+1)}\right)\right\} \geq 1-\alpha.$$

Additionally, if the scores $\left\{S^{(k)}\right\}_{k\in\mathrm{Cal}} \cup \{S_{n+1}\}$ *are a.s. distinct:*

$$\mathbb{P}\left\{Y^{(n+1)} \in \widehat{C}_\alpha\left(X^{(n+1)}\right)\right\} \leq 1-\alpha+\frac{1}{\#\mathrm{Cal}+1}.$$

- ✓ Distribution-free, only requires exchangeability
- ✓ Any regression algorithm (neural nets, random forest...)
- ✓ Finite sample
- ✗ Marginal coverage: $\mathbb{P}\left\{Y^{(n+1)} \in \widehat{C}_\alpha\left(X^{(n+1)}\right) \cancel{| X^{(n+1)} = x}\right\} \geq 1-\alpha$

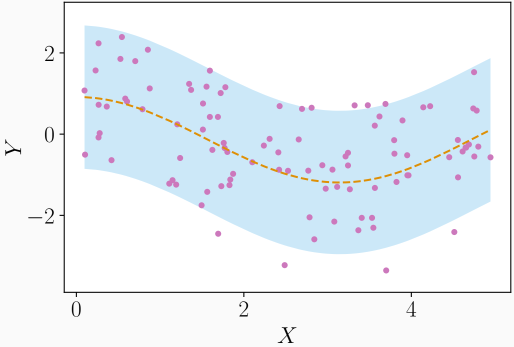
(Way too short) Intro to (Split) Conformal Prediction

Standard Split Conformal Prediction for Mean-Regression

Improving Adaptiveness: Conformalized Quantile Regression

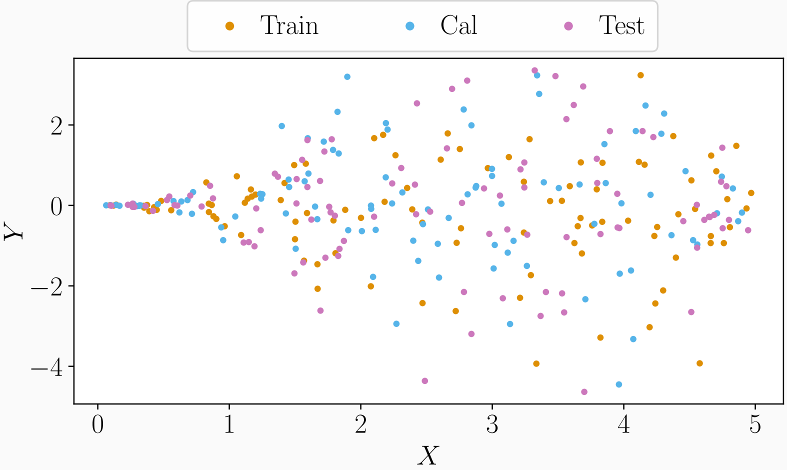
Conformal Prediction with Missing Values

Standard mean-regression SCP is not adaptive



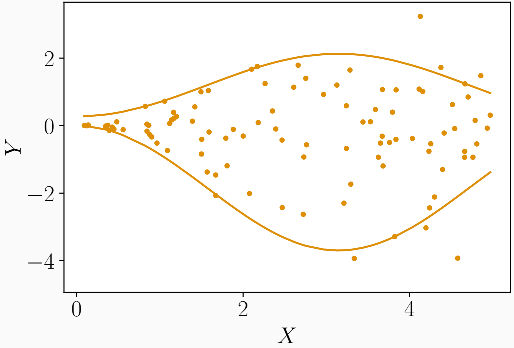
- Predict with $\hat{\mu}$
- Build $\widehat{C}_\alpha(x)$: $[\hat{\mu}(x) \pm q_{1-\alpha}(\mathcal{S})]$

Conformalized Quantile Regression (CQR)[4]



[4] Romano et al. (2019), *Conformalized Quantile Regression*, NeurIPS

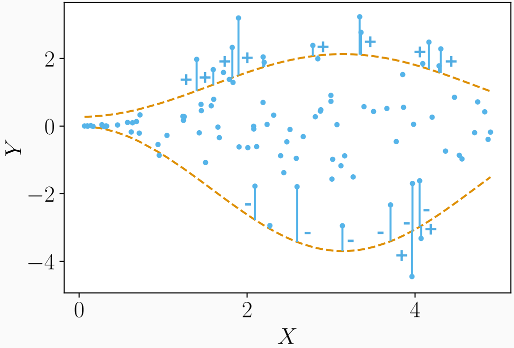
Conformalized Quantile Regression (CQR)[4]: training step



- Learn (or get) $\widehat{QR}_{\text{lower}}$ and $\widehat{QR}_{\text{upper}}$

[4] Romano et al. (2019), *Conformalized Quantile Regression*, NeurIPS

Conformalized Quantile Regression (CQR)[4]: calibration step

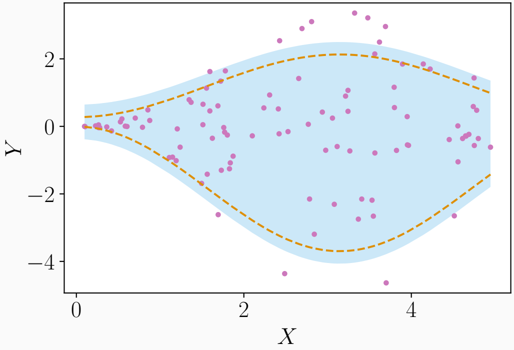


- Predict with $\widehat{QR}_{\text{lower}}$ and $\widehat{QR}_{\text{upper}}$
- Get the scores $\mathcal{S} = \left\{ S^{(k)} \right\}_{\mathrm{Cal}} \cup \{+\infty\}$
- Compute the $(1-\alpha)$ empirical quantile of $\mathcal{S}$, noted $q_{1-\alpha}(\mathcal{S})$

$$\hookrightarrow \quad S^{(k)} := \max\left\{ \widehat{QR}_{\text{lower}}\left(X^{(k)}\right) - Y^{(k)}, Y^{(k)} - \widehat{QR}_{\text{upper}}\left(X^{(k)}\right) \right\}$$

[4] Romano et al. (2019), *Conformalized Quantile Regression*, NeurIPS

Conformalized Quantile Regression (CQR)[4]: prediction step



- Predict with $\widehat{QR}_{\text{lower}}$ and $\widehat{QR}_{\text{upper}}$

- Build

$$\widehat{C}_\alpha(x) = [\widehat{QR}_{\text{lower}}(x) - q_{1-\alpha}(\mathcal{S}); \widehat{QR}_{\text{upper}}(x) + q_{1-\alpha}(\mathcal{S})]$$

[4] Romano et al. (2019), *Conformalized Quantile Regression*, NeurIPS

CQR: theoretical guarantees

CQR enjoys finite sample guarantees proved in Romano et al. (2019), as a particular case of Split Conformal Prediction (SCP).

Theorem

Suppose $\left(X^{(k)}, Y^{(k)}\right)_{k=1}^{n+1}$ *are exchangeable (or i.i.d.). CQR applied on* $\left(X^{(k)}, Y^{(k)}\right)_{k=1}^{n}$ *outputs* $\widehat{C}_\alpha(\cdot)$ *such that:*

$$\mathbb{P}\left\{Y^{(n+1)} \in \widehat{C}_\alpha\left(X^{(n+1)}\right)\right\} \geq 1 - \alpha.$$

Additionally, if the scores $\left\{S^{(k)}\right\}_{k \in \mathrm{Cal}} \cup \{S_{n+1}\}$ *are a.s. distinct:*

$$\mathbb{P}\left\{Y^{(n+1)} \in \widehat{C}_\alpha\left(X^{(n+1)}\right)\right\} \leq 1 - \alpha + \frac{1}{\#\mathrm{Cal} + 1}.$$

- ✓ Distribution-free, only requires exchangeability
- ✓ Any quantile regression algorithm (neural nets, random forest...)
- ✓ Finite sample
- ✗ Marginal coverage: $\mathbb{P}\left\{Y^{(n+1)} \in \widehat{C}_\alpha\left(X^{(n+1)}\right) \cancel{\mid X^{(n+1)} = x}\right\} \geq 1 - \alpha$ conditional

(Way too short) Intro to (Split) Conformal Prediction

Conformal Prediction with Missing Values

Missing values are ubiquitous and challenging

Data: $\left(X^{(k)}, M^{(k)}, Y^{(k)}\right)_{k=1}^{n}$

Y	X_1	X_2	X_3
22.42	0.55	0.67	0.03
8.26	0.72	0.18	0.55
19.41	0.60	0.58	NA
19.75	0.54	0.43	0.96
7.32	NA	0.19	NA
13.55	0.65	0.69	0.50
20.75	NA	NA	0.61
9.26	0.89	NA	0.84
9.68	0.963	0.45	0.65

Mask $M =$

$(M_1$	M_2	$M_3)$
0	0	0
0	0	0
0	0	1
0	0	0
1	0	1
0	0	0
1	1	0
0	1	0
0	0	0

$\hookrightarrow$ 2^d potential masks.

$\hookrightarrow$ M can depend on X or Y (depending on the missing mechanism).

$\Rightarrow$ Statistical and computational challenges.

Supervised learning with missing values: impute-then-regress

Impute-then-regress procedures are widely used.

1. Replace `NA` using an imputation function (e.g. the mean), noted ϕ.

$$\begin{array}{c} x^{(1)} \\ x^{(2)} \\ x^{(3)} \\ x^{(4)} \end{array} \begin{bmatrix} -1 & -10 & 6 & 0 \\ 4 & \mathtt{NA} & -2 & 2 \\ 5 & 1 & 2 & \mathtt{NA} \\ 0 & \mathtt{NA} & \mathtt{NA} & 1 \end{bmatrix} \xrightarrow{\phi} \begin{array}{c} u^{(1)} \\ u^{(2)} \\ u^{(3)} \\ u^{(4)} \end{array} \begin{bmatrix} -1 & -10 & 6 & 0 \\ 4 & -4.5 & -2 & 2 \\ 5 & 1 & 2 & 1 \\ 0 & -4.5 & 3 & 1 \end{bmatrix}$$

2. Train your algorithm (Random Forest, Neural Nets, etc.) on the imputed data:

$$\left\{ \underbrace{\phi\left(X^{(k)}, M^{(k)}\right)}_{U^{(k)} = \text{imputed } X^{(k)}}, Y^{(k)} \right\}_{k=1}^{n}.$$

$\hookrightarrow$ we consider an impute-then-regress pipeline in this work.

Predictive uncertainty quantification with missing values

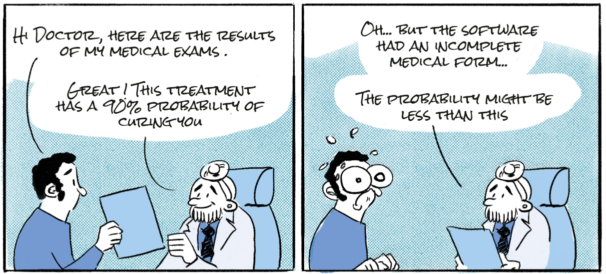
Goal: predict $Y^{(n+1)}$ with **confidence** $1-\alpha$, i.e. build the smallest $\mathcal{C}_\alpha$ such that:

1. Marginal Validity (MV)

$$\mathbb{P}\left\{Y^{(n+1)} \in \mathcal{C}_\alpha\left(X^{(n+1)}, M^{(n+1)}\right)\right\} \geq 1-\alpha. \qquad \text{(MV)}$$

2. Mask-Conditional-Validity (MCV)

$$\forall m \in \{0,1\}^d : \mathbb{P}\left\{Y^{(n+1)} \in \mathcal{C}_\alpha\left(X^{(n+1)}, m\right) | M^{(n+1)} = m\right\} \geq 1-\alpha. \qquad \text{(MCV)}$$



Illustrations @theoremlinger

CP is marginally valid (MV) after imputation

To apply conformal prediction we need **exchangeable** data.

Lemma

Assume $\left(X^{(k)}, M^{(k)}, Y^{(k)}\right)_{k=1}^{n}$ *are i.i.d. (or exchangeable).*

Then, for any missing mechanism, for almost all imputation function[5] ϕ*:*

$\left(\phi\left(X^{(k)}, M^{(k)}\right), Y^{(k)}\right)_{k=1}^{n}$ *are* **exchangeable**.

$\Rightarrow$ CQR, and Conformal Prediction, applied on an imputed data set still enjoys marginal guarantees[6]:

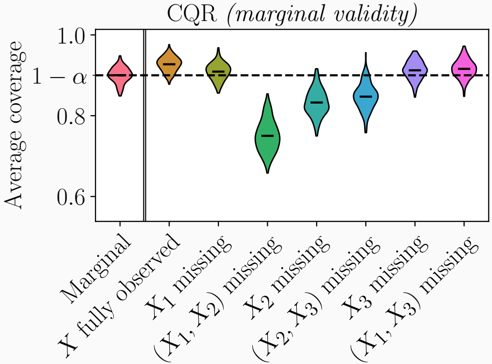
$$\mathbb{P}\left\{Y^{(n+1)} \in \widehat{C}_\alpha\left(X^{(n+1)}, M^{(n+1)}\right)\right\} \geq 1 - \alpha.$$

[5] Even if the imputation is not accurate, the guarantee will hold.

[6] The upper bound also holds under continuously distributed scores.

CQR is marginally valid on imputed data sets

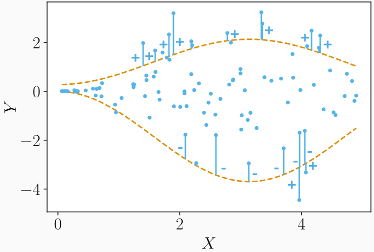
$$Y = \beta^T X + \varepsilon,\ \beta = (1, 2, -1)^T,\ X \text{ and } \varepsilon \text{ Gaussian.}$$



- The predictive uncertainty strongly depends on the mask

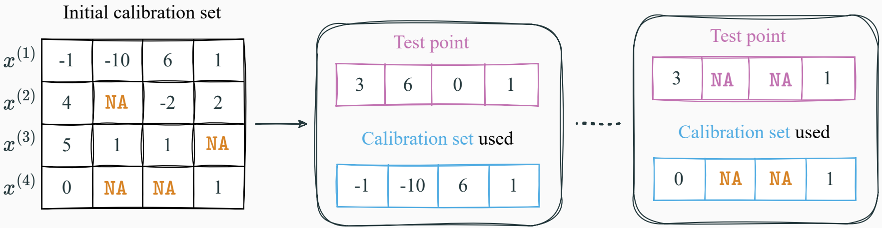
	Imputation+CQR
(MV)	✓
(MCV)	✗

Observation: the α-correction term is computed among all the data points, regardless of their mask!



Warning: 2^d possible masks

$\Rightarrow$ Splitting the calibration set by mask is infeasible (lack of data)!



Missing Data Augmentation (MDA)

Idea: for each test point, modify the calibration points to mimic the test mask

Test point

3	NA	NA	1

Initial calibration set

$x^{(1)}$	-1	-10	6	1
$x^{(2)}$	4	NA	-2	2
$x^{(3)}$	5	1	1	NA
$x^{(4)}$	0	NA	NA	1

Calibration set used

$\tilde{x}^{(1)}$	-1	NA	NA	1
$\tilde{x}^{(2)}$	4	NA	NA	2
$\tilde{x}^{(3)}$	5	NA	NA	NA
$\tilde{x}^{(4)}$	0	NA	NA	1

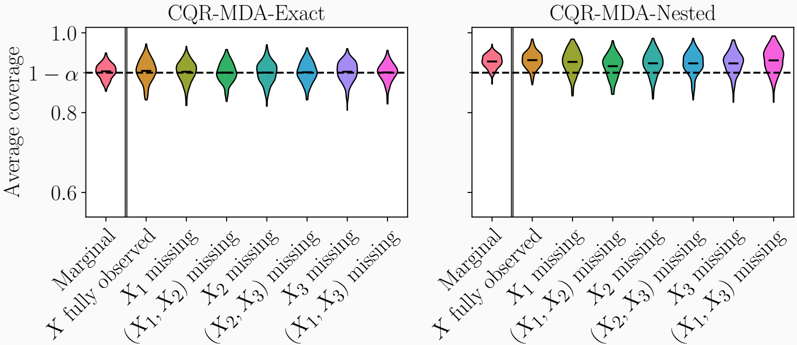
Algorithms: MDA with Exact masking or with Nested masking.

CP-MDA-Exact

MDA achieves Mask-Conditional-Validity (MCV)

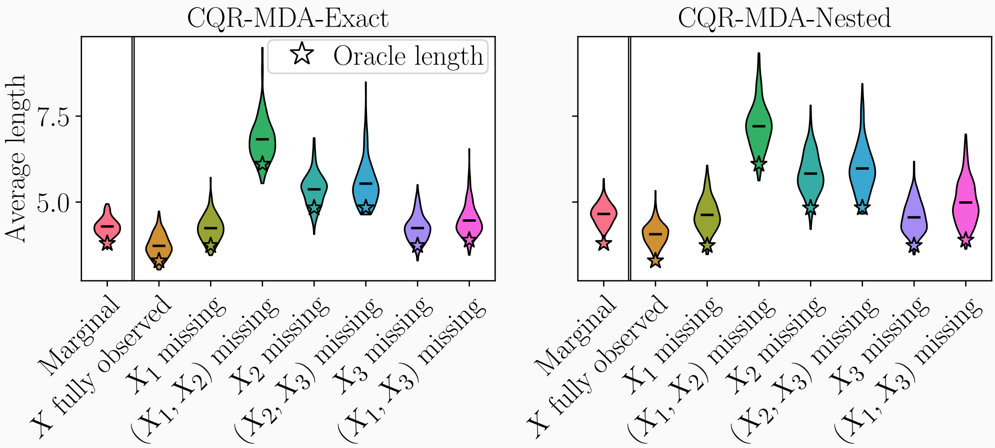
Theorem (Informal)

If $M \perp\!\!\!\perp (X, Y)$, for almost all imputation function, CP-MDA reaches (MCV).



	Imputation+CQR	CQR-MDA
(MV)	✓	✓
(MCV)	✗	✓

MDA achieves (MCV) in an informative way



Take-home-messages

- CP marginal guarantees hold on the imputed data set.
- Missingness introduces additional heteroskedasticity, creating a need for quantile regression based methods.
- CQR fails to attain coverage conditional on the missing pattern.
- **Missing data augmentation is the first method to output predictive intervals with missing values**.
- Missing data augmentation attains conditional coverage with respect to the missing pattern (in MCAR setting).
- Extensions:
    - Synthetic experiments in higher dimension
    - Semi-synthetic experiments
    - Synthetic experiments beyond MCAR (MAR and MNAR)
    - Real data experiments (TraumaBase)
    - CP-MDA-Nested link to CP-MDA-Nested, an algorithm which does not discard any calibration point
    - Consistency of universal quantile learner when chained with almost any

Questions? :)

Thanks for listening and feel free to reach out!

Paper ⟶

Code ⟶

Summary ⟶

References i

Barber, R. F., Candès, E. J., Ramdas, A., and Tibshirani, R. J. (2021). The limits of distribution-free conditional predictive inference. *Information and Inference: A Journal of the IMA*, 10(2).

Chernozhukov, V., Wüthrich, K., and Zhu, Y. (2021). Distributional conformal prediction. *Proceedings of the National Academy of Sciences*, 118(48).

Gibbs, I., Cherian, J. J., and Candès, E. J. (2023). Conformal prediction with conditional guarantees. arXiv: 2305.12616.

Guan, L. (2022). Localized conformal prediction: a generalized inference framework for conformal prediction. *Biometrika*, 110(1).

Izbicki, R., Shimizu, G., and Stern, R. B. (2022). CD-split and HPD-split: Efficient conformal regions in high dimensions. *Journal of Machine Learning Research*, 23(87).

Jung, C., Noarov, G., Ramalingam, R., and Roth, A. (2023). Batch multivalid conformal prediction. In *International Conference on Learning Representations*.

Kivaranovic, D., Johnson, K. D., and Leeb, H. (2020). Adaptive, Distribution-Free Prediction Intervals for Deep Networks. In *International Conference on Artificial Intelligence and Statistics*. PMLR.

Lei, J., G'Sell, M., Rinaldo, A., Tibshirani, R. J., and Wasserman, L. (2018). Distribution-Free Predictive Inference for Regression. *Journal of the American Statistical Association*.

Lei, J. and Wasserman, L. (2014). Distribution-free prediction bands for non-parametric regression. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 76(1).

Papadopoulos, H., Proedrou, K., Vovk, V., and Gammerman, A. (2002). Inductive Confidence Machines for Regression. In *Machine Learning: ECML*. Springer.

Romano, Y., Barber, R. F., Sabatti, C., and Candès, E. (2020). With Malice Toward None: Assessing Uncertainty via Equalized Coverage. *Harvard Data Science Review*, 2(2).

Romano, Y., Patterson, E., and Candès, E. (2019). Conformalized Quantile Regression. In *Advances in Neural Information Processing Systems*. Curran Associates, Inc.

Sesia, M. and Romano, Y. (2021). Conformal prediction using conditional histograms. In *Advances in Neural Information Processing Systems*. Curran Associates, Inc.

Vovk, V. (2012). Conditional Validity of Inductive Conformal Predictors. In *Asian Conference on Machine Learning*. PMLR.

Vovk, V., Gammerman, A., and Shafer, G. (2005). *Algorithmic Learning in a Random World*. Springer US.

Informative conditional coverage as such is impossible

- Impossibility results

  $\hookrightarrow$ Lei and Wasserman (2014); Vovk (2012); Barber et al. (2021)

> Without distribution assumption, in finite sample, a perfectly conditionally valid $\widehat{C}_\alpha$ is such that $\mathbb{P}\left\{\text{mes}\left(\widehat{C}_\alpha(x)\right) = \infty\right\} = 1$ for any non-atomic x.

- Approximate conditional coverage

  $\hookrightarrow$ Romano et al. (2020); Guan (2022); Jung et al. (2023); Gibbs et al. (2023)

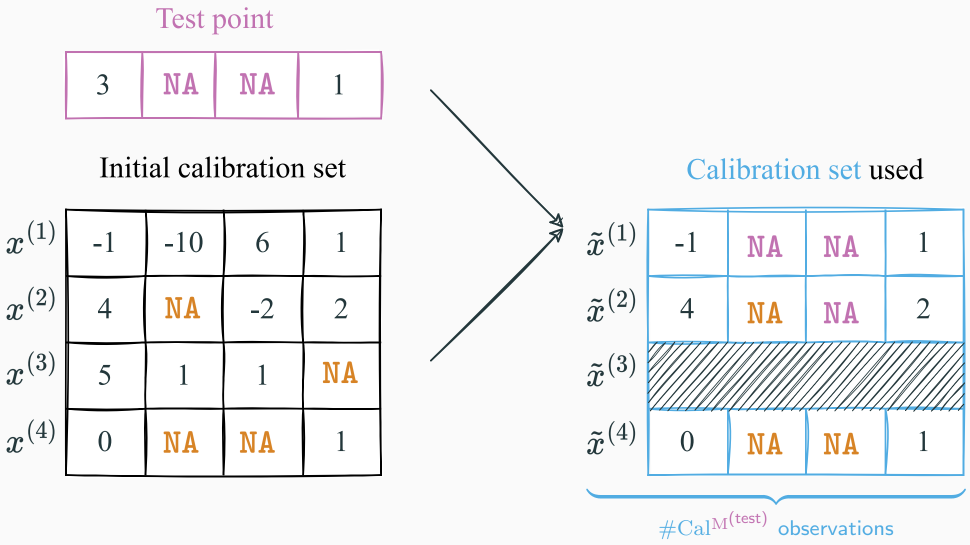
  Target $\mathbb{P}(Y_{n+1} \in \widehat{C}_\alpha | X_{n+1} \in \mathcal{R}(x)) \geq 1 - \alpha$

- Asymptotic (with the sample size) conditional coverage

  $\hookrightarrow$ Romano et al. (2019); Kivaranovic et al. (2020); Chernozhukov et al. (2021); Sesia and Romano (2021); Izbicki et al. (2022)

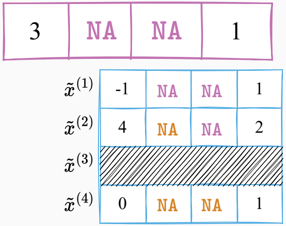
Non exhaustive references.

CP-MDA with Exact masking



CQR-MDA with exact masking in words

1. Split the training set into a proper training set and calibration set
2. Train the imputation function on the proper training set
3. Impute the proper training set
4. Train the quantile regressors on the imputed proper training set
5. For a test point $\left(X^{(n+1)}, M^{(n+1)}\right)$:



   5.1 For each $j \in [\![1, d]\!]$ s.t. $M_j^{(n+1)} = 1$, set $\tilde{M}_j^{(k)} = 1$ for k in Cal s.t. $M^{(k)} \subset M^{(n+1)}$
   5.2 Impute the new calibration set
   5.3 Compute the calibration correction, i.e. $q_{1-\alpha}(\mathcal{S})$
   5.4 Impute the test point
   5.5 Predict with the quantile regressors and the correction previously obtained, $q_{1-\alpha}(\mathcal{S})$

CP-MDA-Nested

CP-MDA-Exact reminder

Test point

3	NA	NA	1

Initial calibration set

$x^{(1)}$	-1	-10	6	1
$x^{(2)}$	4	NA	-2	2
$x^{(3)}$	5	1	1	NA
$x^{(4)}$	0	NA	NA	1

Calibration set used

$\tilde{x}^{(1)}$	-1	NA	NA	1
$\tilde{x}^{(2)}$	4	NA	NA	2
$\tilde{x}^{(3)}$				
$\tilde{x}^{(4)}$	0	NA	NA	1

What if we kept all individuals?

Test point

3	NA	NA	1

Initial calibration set

$x^{(1)}$	-1	-10	6	1
$x^{(2)}$	4	NA	-2	2
$x^{(3)}$	5	1	1	NA
$x^{(4)}$	0	NA	NA	1

Calibration set used

$\tilde{x}^{(1)}$	-1	NA	NA	1
$\tilde{x}^{(2)}$	4	NA	NA	2
$\tilde{x}^{(3)}$	5	NA	NA	NA
$\tilde{x}^{(4)}$	0	NA	NA	1

Idea: modify the test point accordingly

Test point

3	NA	NA	1

Initial calibration set

$x^{(1)}$	-1	-10	6	1
$x^{(2)}$	4	NA	-2	2
$x^{(3)}$	5	1	1	NA
$x^{(4)}$	0	NA	NA	1

Calibration set used

$\tilde{x}^{(1)}$	-1	NA	NA	1
$\tilde{x}^{(2)}$	4	NA	NA	2
$\tilde{x}^{(3)}$	5	NA	NA	NA
$\tilde{x}^{(4)}$	0	NA	NA	1

and

Temporary test points

3	NA	NA	1
3	NA	NA	1
3	NA	NA	NA
3	NA	NA	1

CQR-MDA with nested masking in words

1. For a test point $(X^{(n+1)}, M^{(n+1)})$:

3	NA	NA	1

$\tilde{x}^{(1)}$	-1	NA	NA	1
$\tilde{x}^{(2)}$	4	NA	NA	2
$\tilde{x}^{(3)}$	5	NA	NA	NA
$\tilde{x}^{(4)}$	0	NA	NA	1

   1.1 Set $\tilde{M}^{(k)} = \max(M^{(k)}, M^{(n+1)})$ for k in the `calibration set`

   1.2 Impute the new `calibration set`

   1.3 For each augmented calibration point k:

      1.3.1 Get its score $S^{(k)}$

      1.3.2 Impute-then-predict on the augmented test point $(X^{(n+1)}, \tilde{M}^{(k)})$, giving: $\widehat{QR}_{\alpha/2}(\tilde{X}^{(n+1),k})$ and $\widehat{QR}_{1-\alpha/2}(\tilde{X}^{(n+1),k})$

3	NA	NA	1
3	NA	NA	1
3	NA	NA	NA
3	NA	NA	1

      1.3.3 Compute the corrected prediction interval:
$$[\widehat{QR}_{\alpha/2}(\tilde{X}^{(n+1),k}) - S^{(k)}; \widehat{QR}_{1-\alpha/2}(\tilde{X}^{(n+1),k}) + S^{(k)}] := \left[Z_{\text{inf}}^{(k)}; Z_{\text{sup}}^{(k)}\right]$$

   1.4 Compute the quantiles $q_\alpha(\{Z_{\text{inf}}^{(k)}\}_{k\in\text{Cal}})$ and $q_{1-\alpha}(\{Z_{\text{sup}}^{(k)}\}_{k\in\text{Cal}})$

   1.5 Predict $[q_\alpha(\{Z_{\text{inf}}^{(k)}\}_{k\in\text{Cal}}); q_{1-\alpha}(\{Z_{\text{sup}}^{(k)}\}_{k\in\text{Cal}})]$

Summary of CP-MDA

Test point

3	NA	NA	1

Initial calibration set

$x^{(1)}$	-1	-10	6	1
$x^{(2)}$	4	NA	-2	2
$x^{(3)}$	5	1	1	NA
$x^{(4)}$	0	NA	NA	1

CP-MDA with exact masking:

calibration set

$\tilde{x}^{(1)}$	-1	NA	NA	1
$\tilde{x}^{(2)}$	4	NA	NA	2
$\tilde{x}^{(3)}$				
$\tilde{x}^{(4)}$	0	NA	NA	1

CP-MDA with nested masking:

calibration set

$\tilde{x}^{(1)}$	-1	NA	NA	1
$\tilde{x}^{(2)}$	4	NA	NA	2
$\tilde{x}^{(3)}$	5	NA	NA	NA
$\tilde{x}^{(4)}$	0	NA	NA	1

and

temporary test points

3	NA	NA	1
3	NA	NA	1
3	NA	NA	NA
3	NA	NA	1