CONFORMAL PREDICTION

MASTER CLASS

Joseba Dalmau

DEEL
Dependable, Explainable & Embeddable Learning

IRT SAINT EXUPÉRY

Uncertainty Quantification in Deep Learning

- Can we trust the predictions of our models?
- Under what circumstances?
- To what extent?

CLASSICAL UQ TECHNIQUES

- Bayesian methods
- Ensembling
- Dropout
- ...

NON POST-HOC

WITHOUT GUARANTEES

Conformal Prediction: Objective and Guarantee

Given: a predictor $\hat{f} : \mathcal{X} \longrightarrow \mathcal{Y}$ and a nominal error rate α.

Build: $\hat{C}_\alpha : \mathcal{X} \longrightarrow \mathcal{P}(\mathcal{Y})$

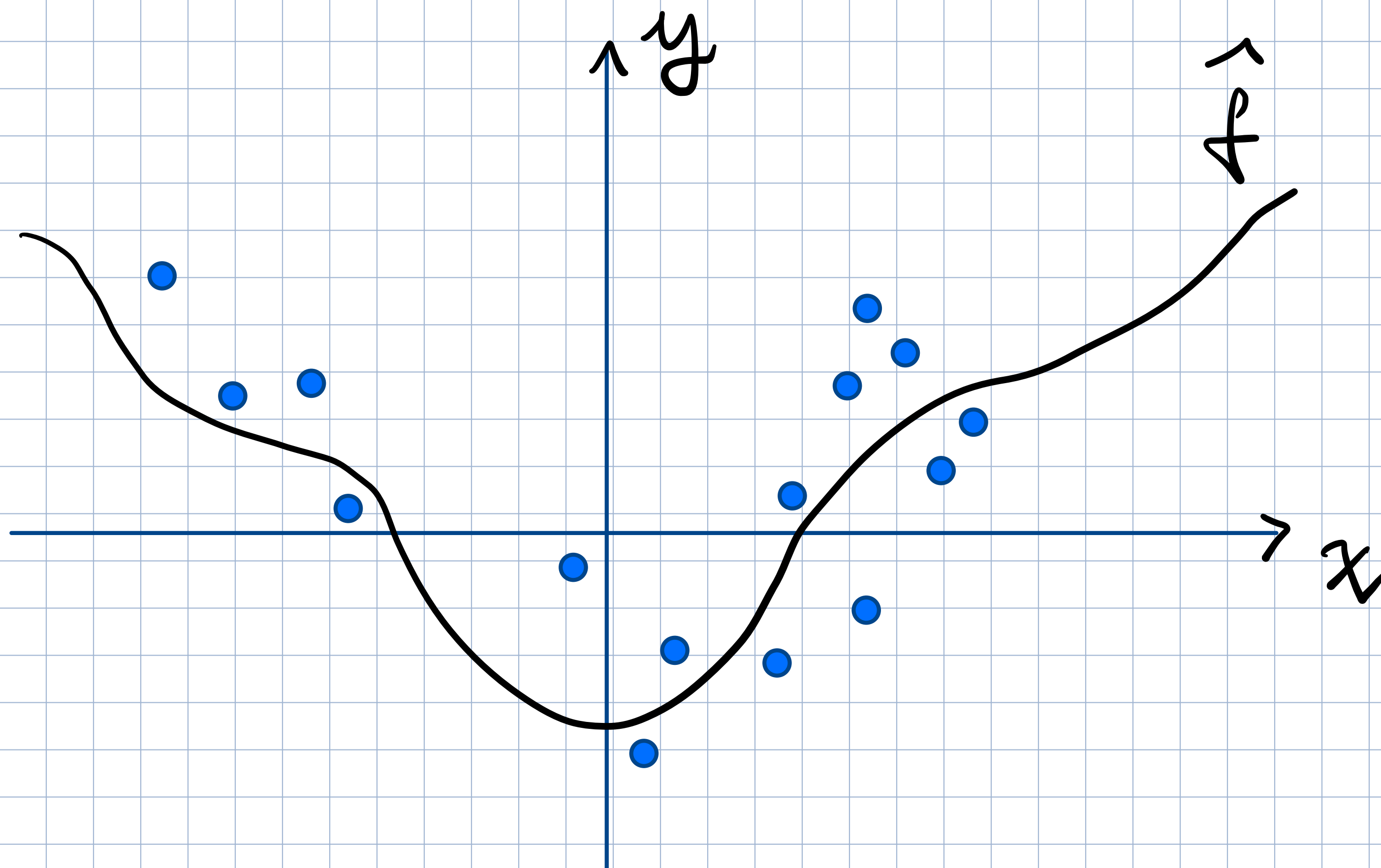
With the following guarantee:

$$\mathbb{P}\left(Y_{test} \in \hat{C}_\alpha(X_{test}) \right) \geq 1 - \alpha$$

CALIBRATION

We use a calibration set

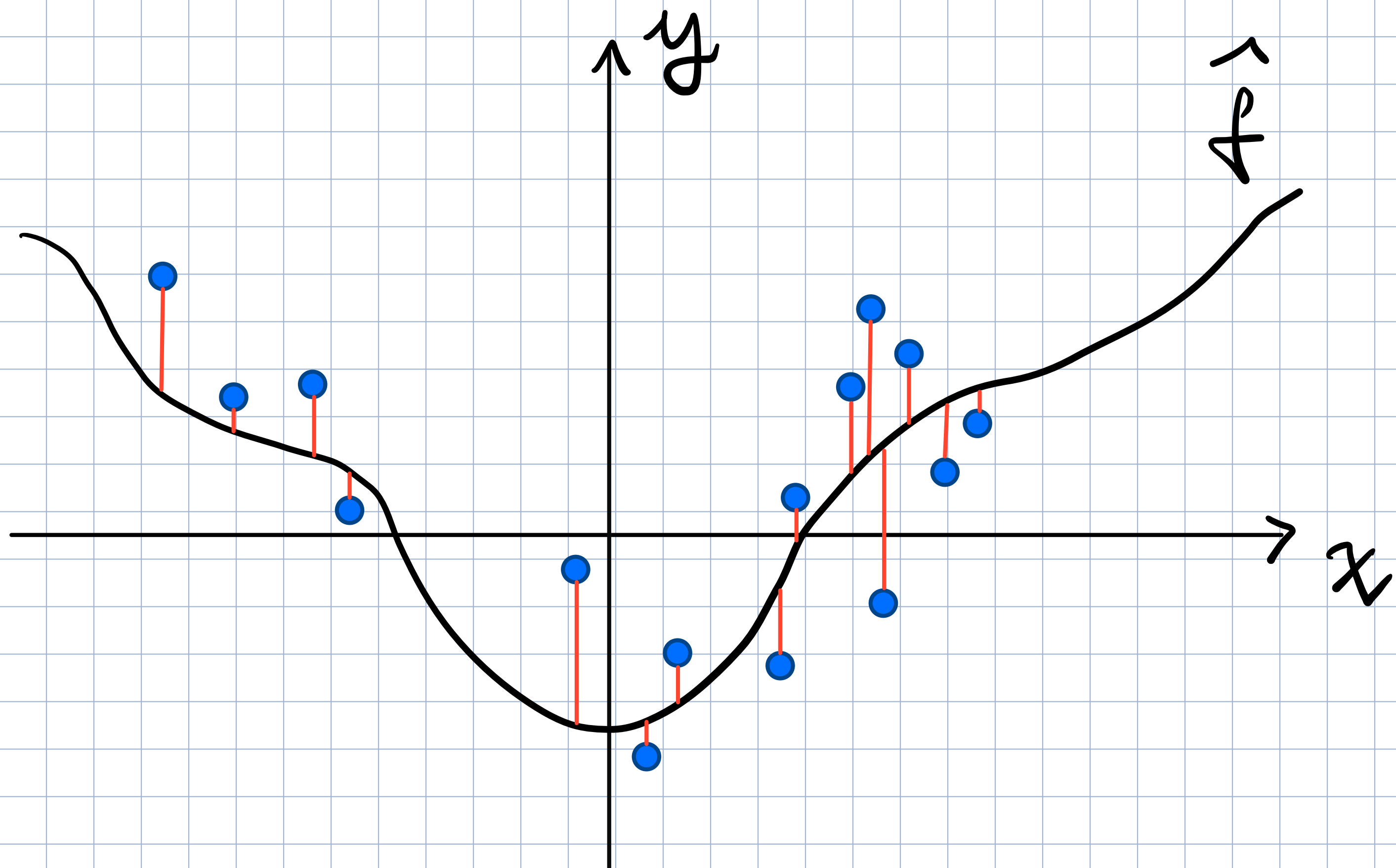
$D_{calib} = \{(X_i, Y_i)\}_{i=1}^{n}$



CALIBRATION

We measure the scores (errors)

$$s_i = |y_i - \hat{f}(x_i)|$$



CALIBRATION

We compute:

$$S_\alpha := \text{the } \frac{\lceil (n+1)(1-\alpha) \rceil}{n}\text{-th}$$

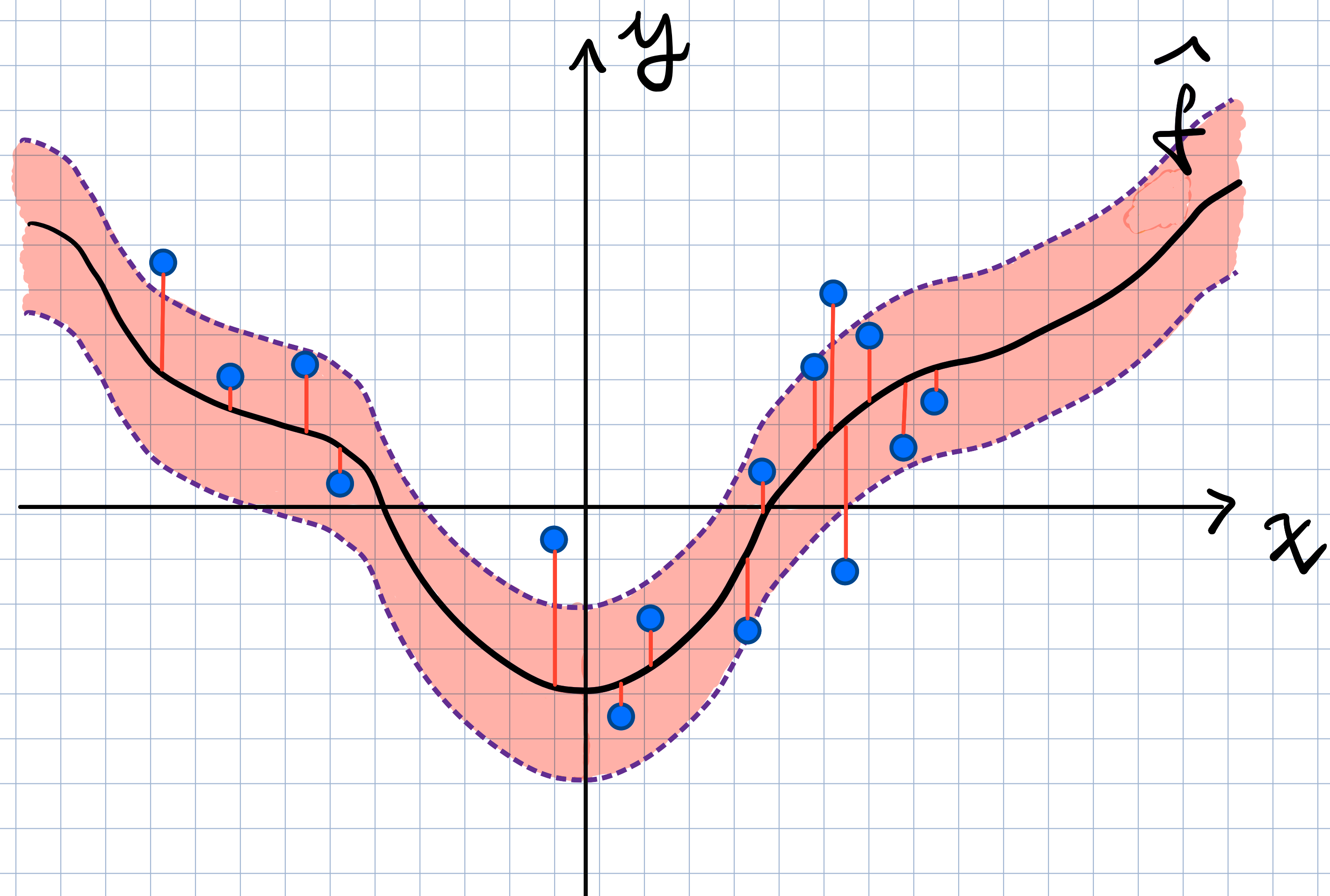
quantile of the scores $S_1, \dots, S_n$

i.e. the $\lceil (n+1)(1-\alpha) \rceil$-th

smallest score.

CALIBRATION

We predict $\hat{C}_\alpha(x) = [\hat{f}(x) - \delta_\alpha, \hat{f}(x) + \delta_\alpha]$

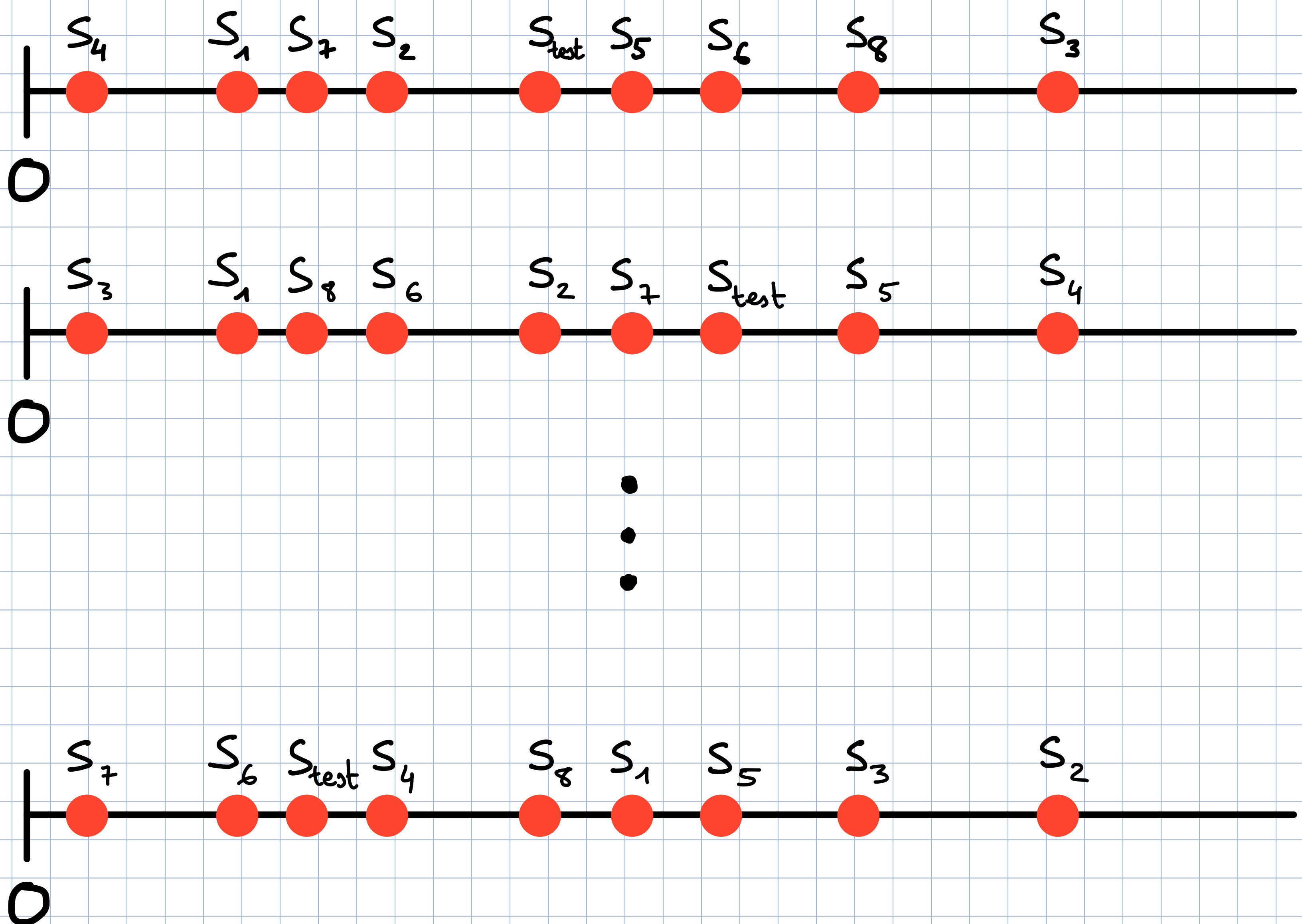


GUARANTEE

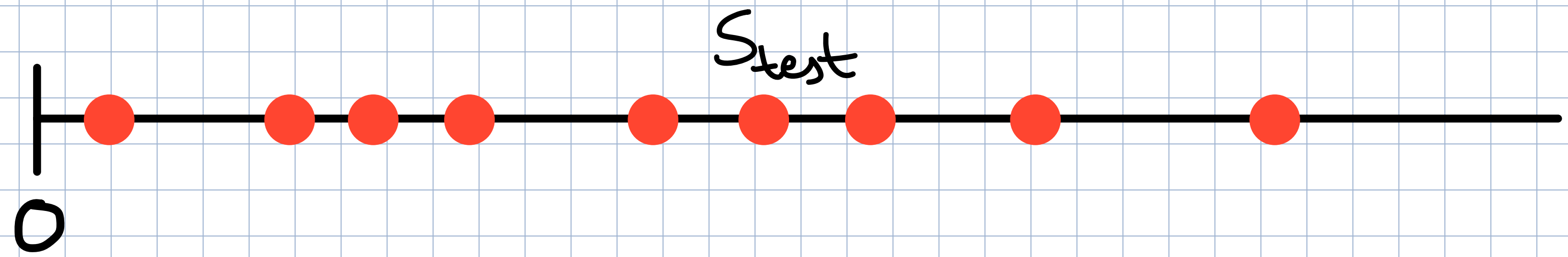
Theorem.– If $\{(X_i, Y_i)\}_{i=1}^{n+1}$ are exchangeable, then:

$$\mathbb{P}\left(Y_{n+1} \in \hat{C}_\alpha(X_{n+1})\right) \geq 1-\alpha$$

PROOF


S4 S1 S7 S2 Stest S5 S6 S8 S3
0
S3 S1 S8 S6 S2 S7 Stest S5 S4
0
S7 S6 Stest S4 S8 S1 S5 S3 S2
0

PROOF



$$\mathbb{P}(\text{Rank of } S_{test} = k)$$

$$= \frac{n!}{(n+1)!} = \frac{1}{n+1}$$

$$\Rightarrow \text{Rank of } S_{test} \sim \text{Unif}(\{1, \dots, n+1\})$$

Proof

Rank of $S_{test} \sim \mathrm{Unif}(\{1, \dots, n+1\})$

$$\Rightarrow \mathbb{P}\left(\begin{matrix}\text{Rank of} \\ S_{test}\end{matrix}\right) = \frac{K}{n+1}$$

We choose the smallest K s.t.

$$\frac{K}{n+1} \geq 1 - \alpha \quad \text{i.e.} \quad K = \lceil (n+1)(1-\alpha) \rceil$$

Too Good To be True?

$\hat{f}$ bad

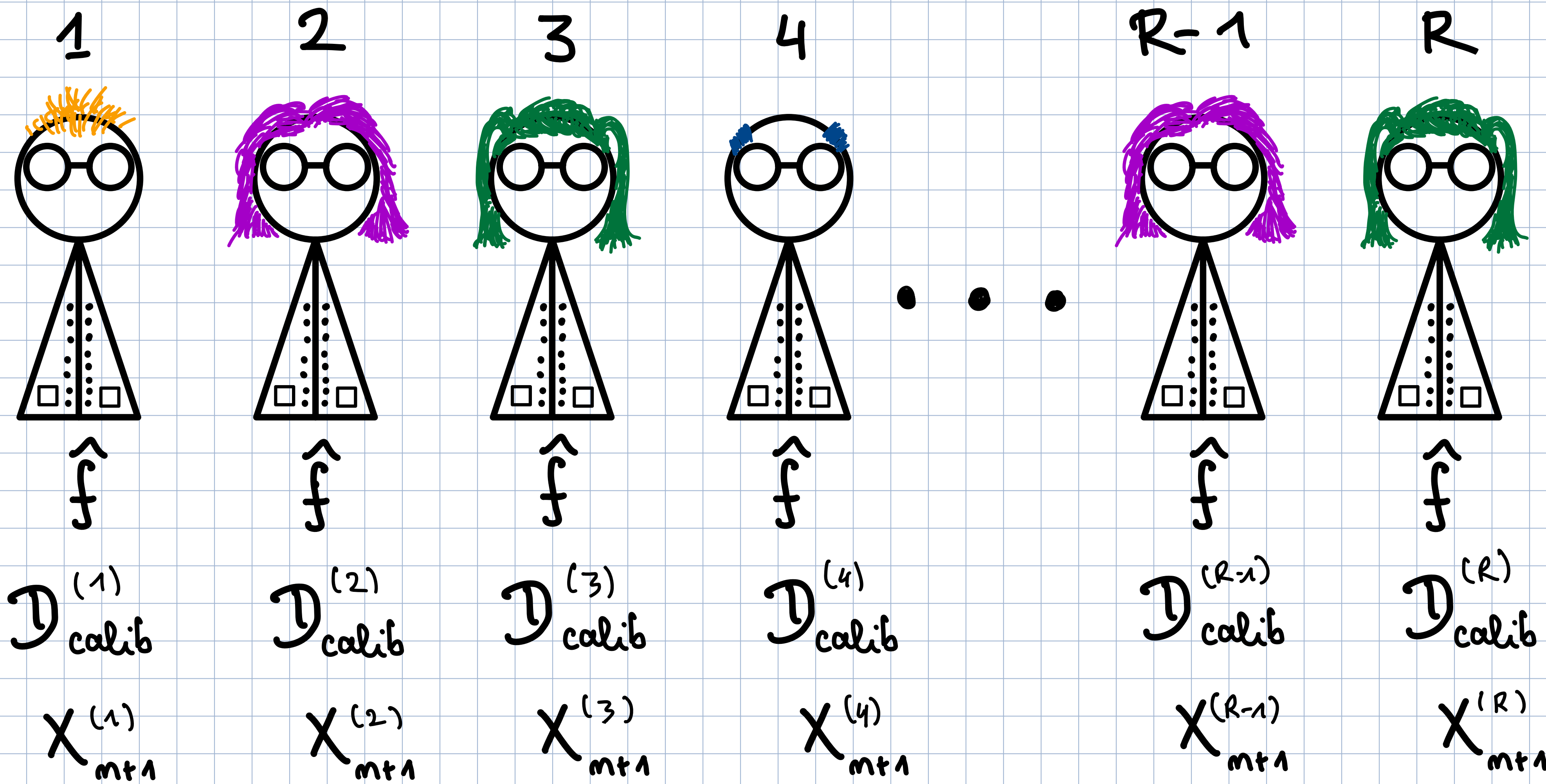
$\Longrightarrow \hat{C}_\alpha$ very large

ADVANTAGES

- Post-hoc
- Distribution-free
- Minimal hypotheses
- Finite-sample guarantee

LIMITATIONS

Calibration-marginal guarantee:

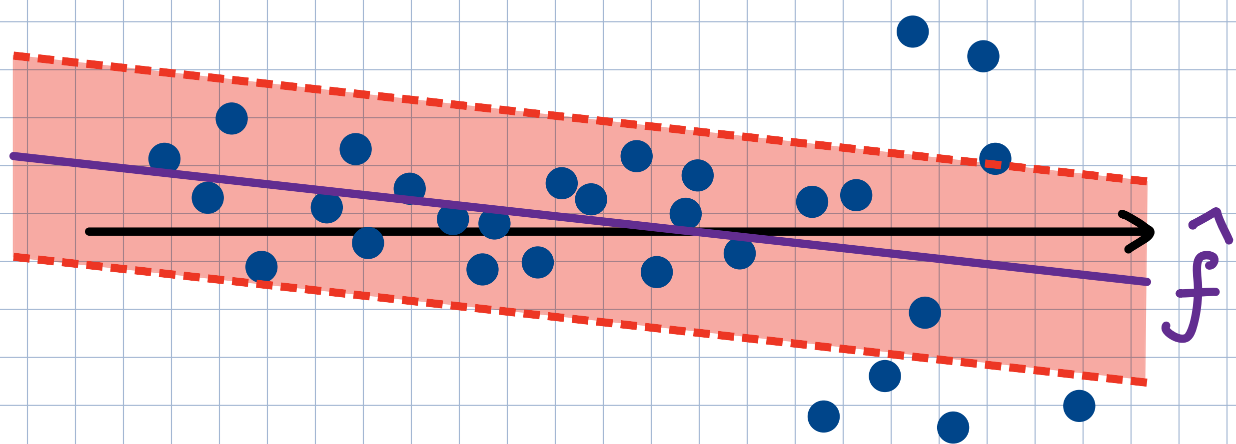


LIMITATIONS

Non-conditional guarantee

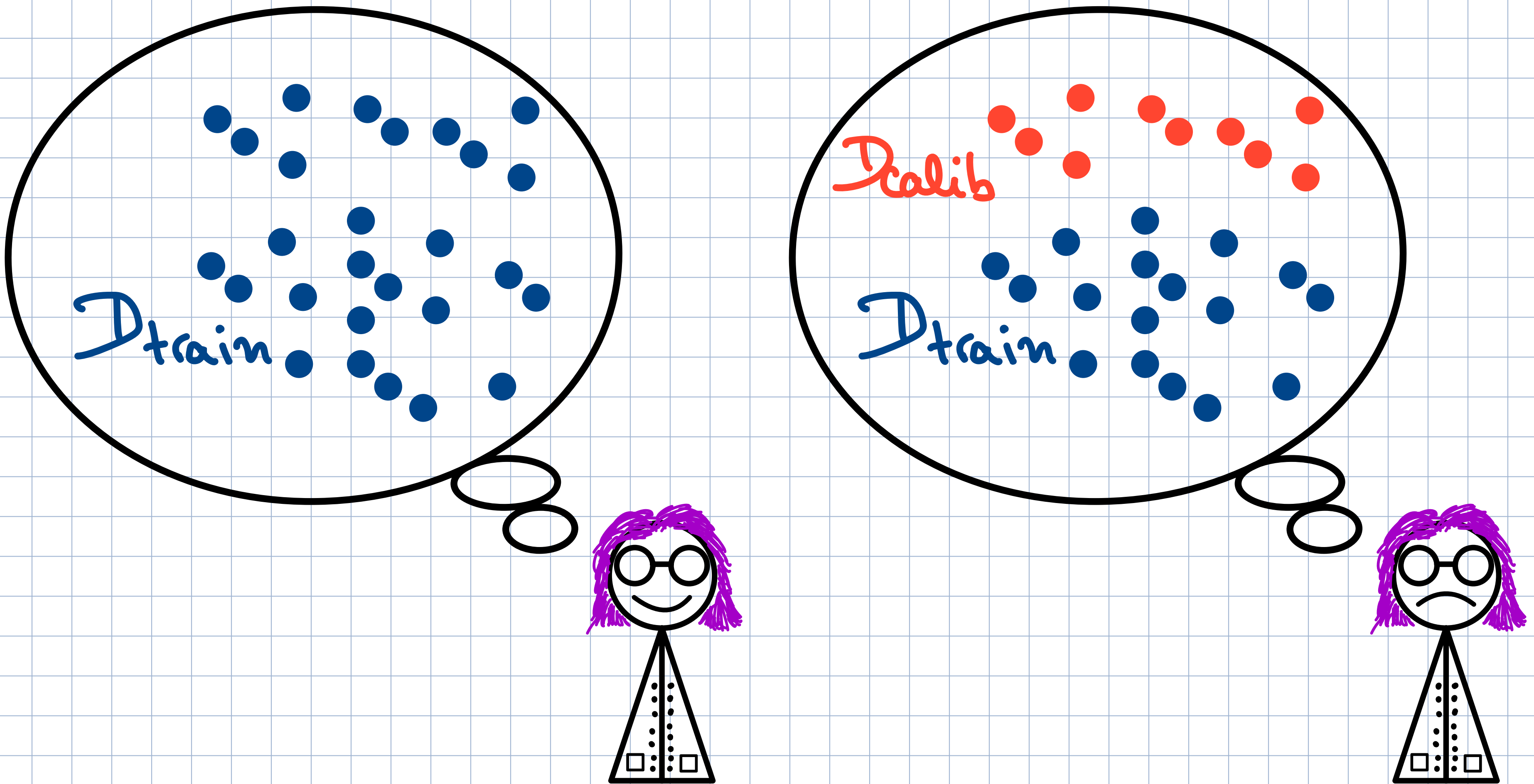
$$\mathbb{P}\left(Y_{n+1} \in \hat{C}_\alpha(X_{n+1})\right) \geq 1-\alpha$$

$$\mathbb{P}\left(Y_{n+1} \in \hat{C}_\alpha(X_{n+1}) \mid X_{n+1} = x\right) \geq 1-\alpha$$



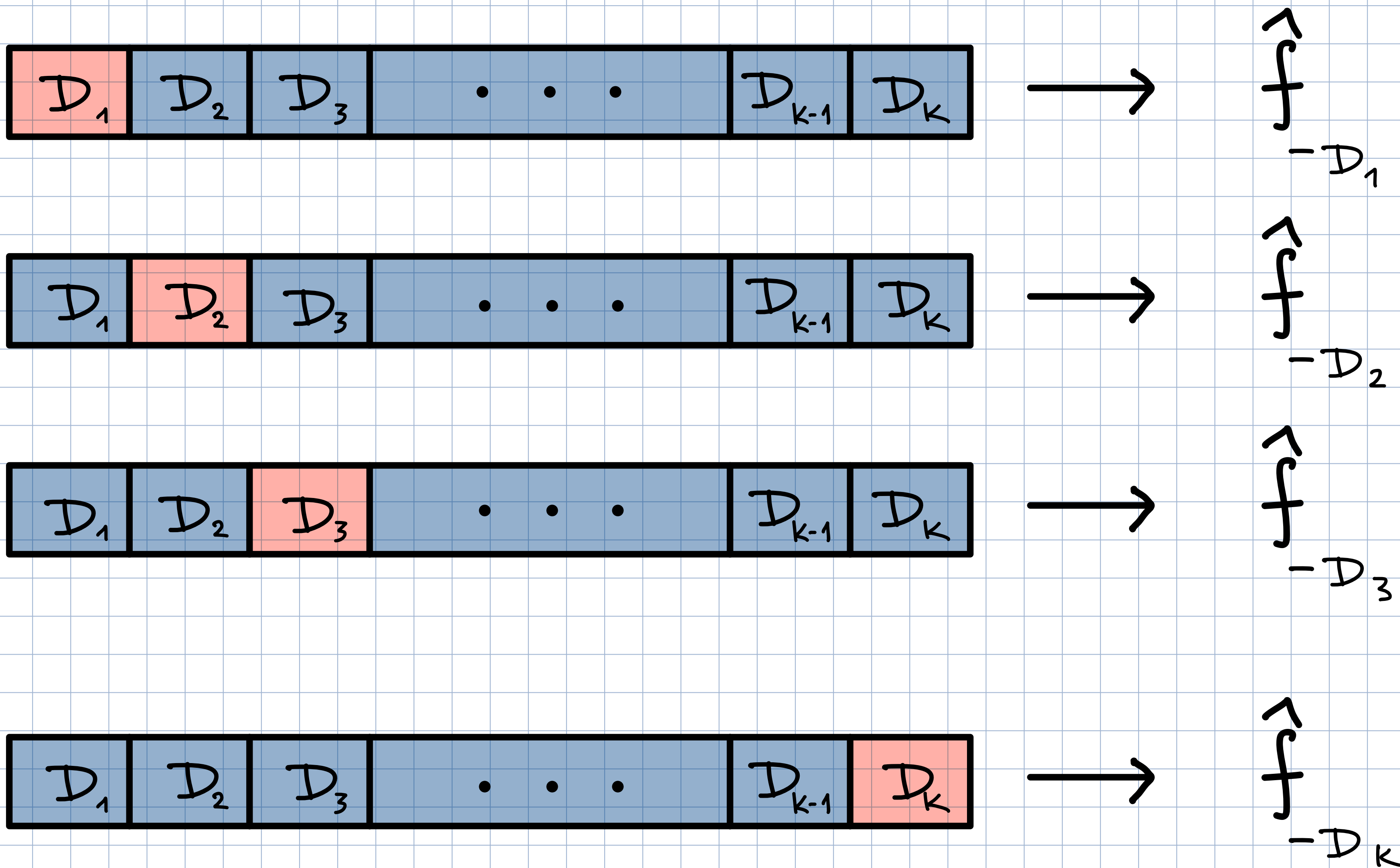
LIMITATIONS

Need for calibration data



JACKNIFE+, CROSS-VALIDATION+

Partition the data:



JACKNIFE+, CROSS-VALIDATION+

Calibration:

$$S_i^{CV} = \left| Y_i - \hat{f}_{-D_{ind(i)}}(X_i) \right|, \quad i = 1, \dots, n$$

Jackknife+, Cross-Validation+

Inference:

$\hat{\ell}_\alpha(x) := \lfloor \alpha(n+1) \rfloor$ - th smallest value of

$$\hat{f}_{-S_{ind(1)}}(x) - S_1^{cv}, \dots, \hat{f}_{-S_{ind(n)}}(x) - S_n^{cv}$$

$\hat{u}_\alpha(x) := \lceil (1-\alpha)(n+1) \rceil$ - th smallest value of

$$\hat{f}_{-S_{ind(1)}}(x) + S_1^{cv}, \dots, \hat{f}_{-S_{ind(n)}}(x) + S_n^{cv}$$

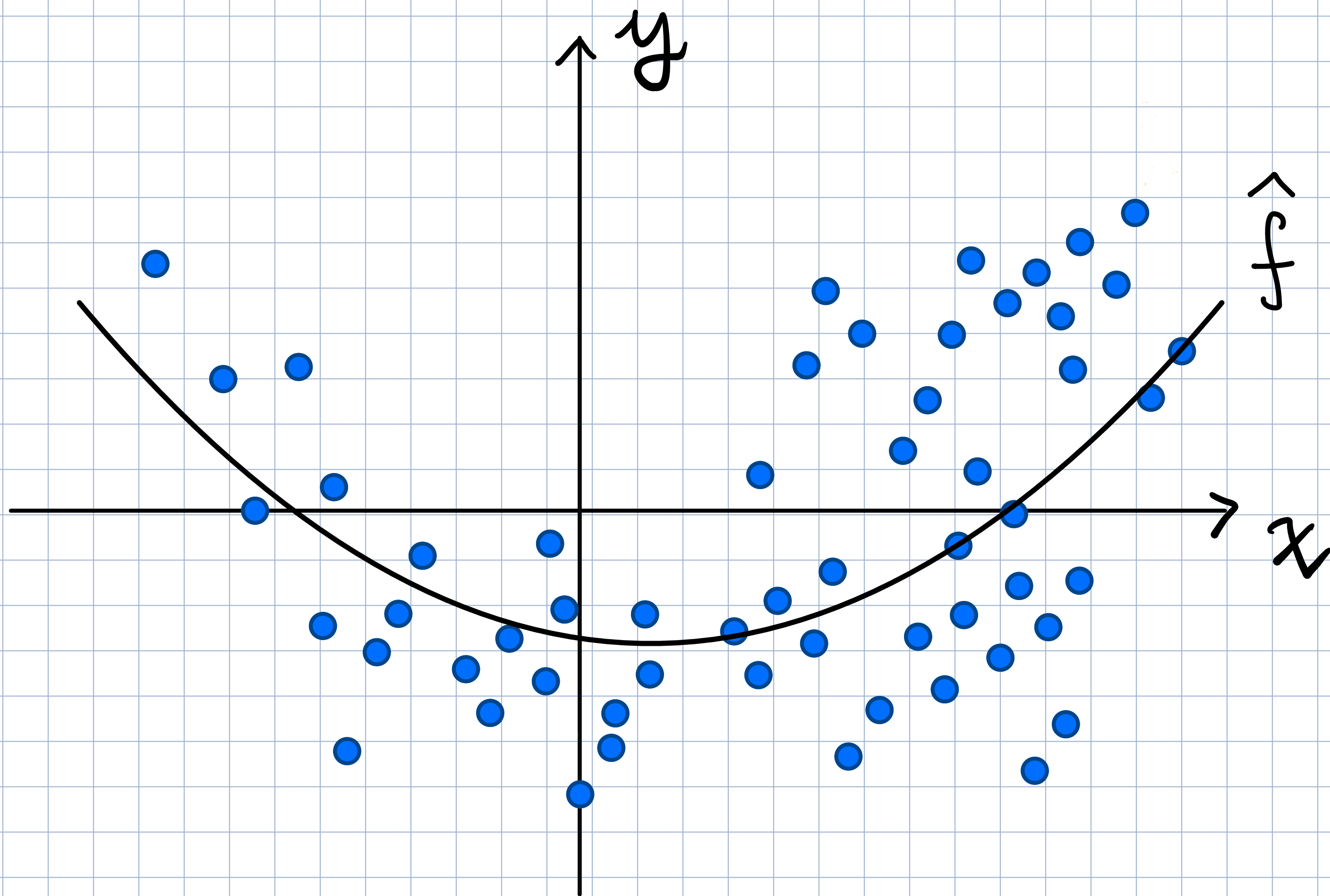
$$\hat{C}_\alpha(x) = [\, \hat{\ell}_\alpha(x) \,,\, \hat{u}(x) \,]$$

JACKNIFE+ , CROSS-VALIDATION+

Guarantee:

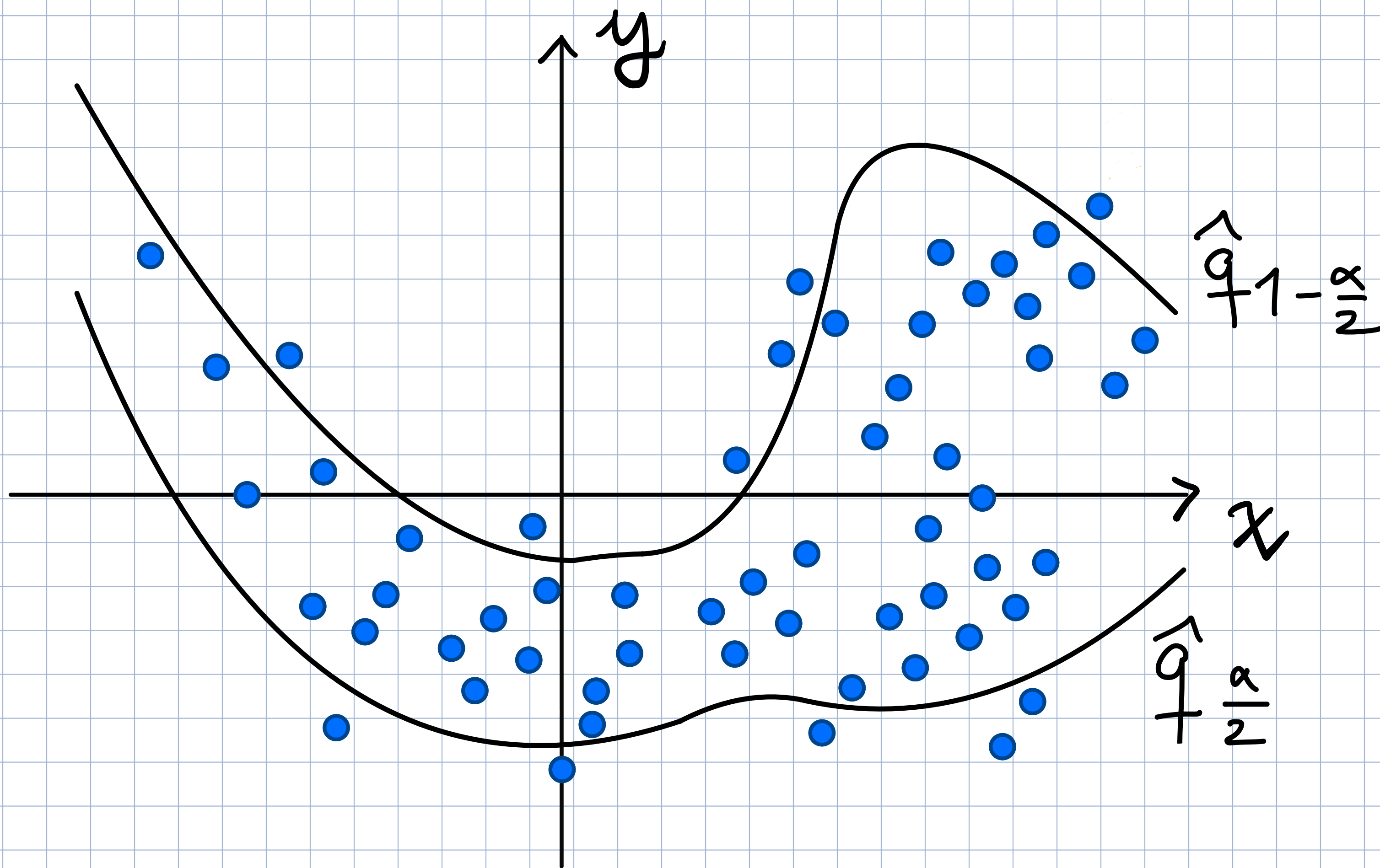
$$\mathbb{P}\left(Y_{n+1} \in \widehat{C}_\alpha(X_{n+1})\right) \geq 1-2\alpha$$

CQR: Conformal Quantile Regression



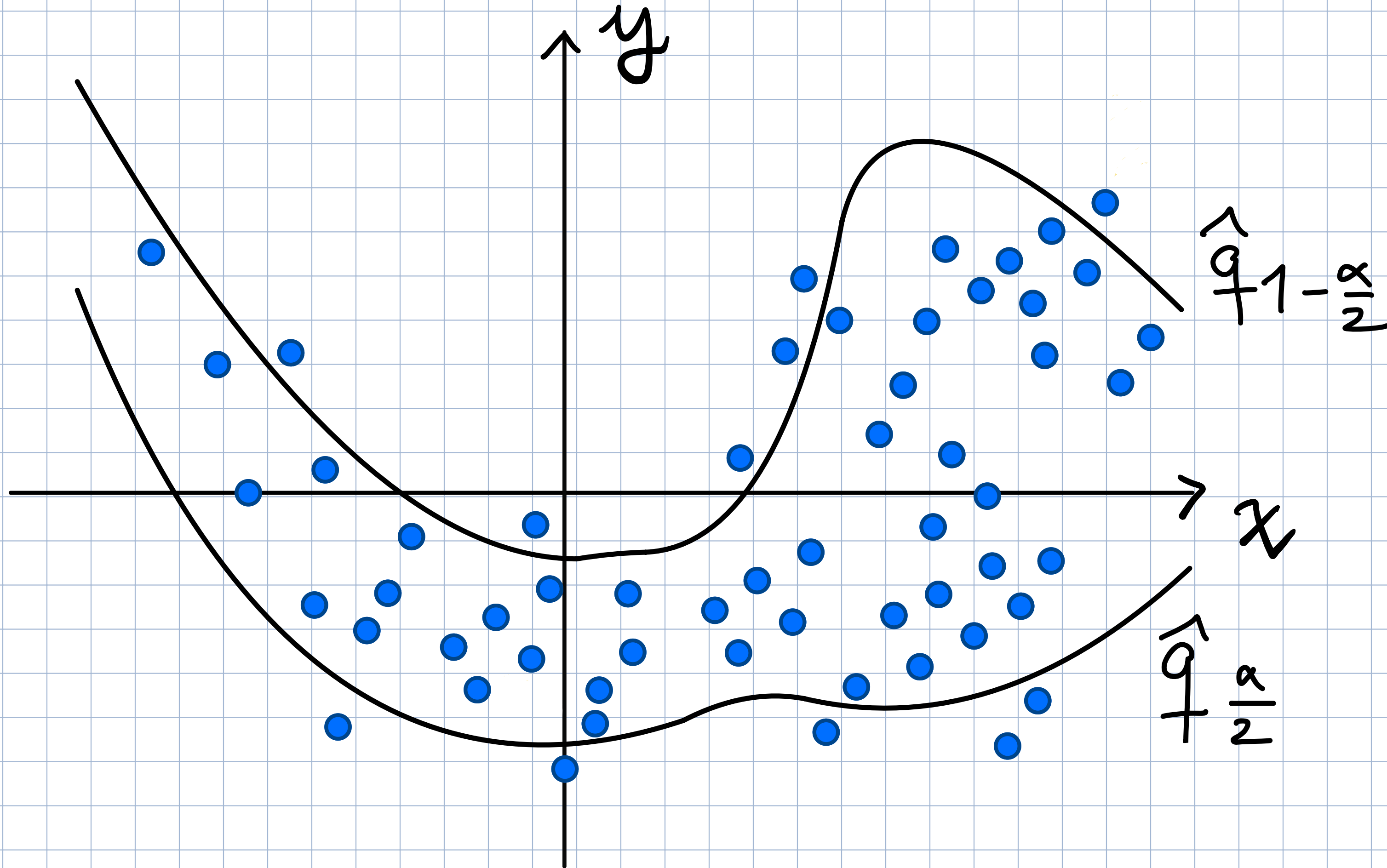
CQR

$$\hat{q}_t(x) \simeq \mathbb{P}(Y \leq t \mid X = x)$$



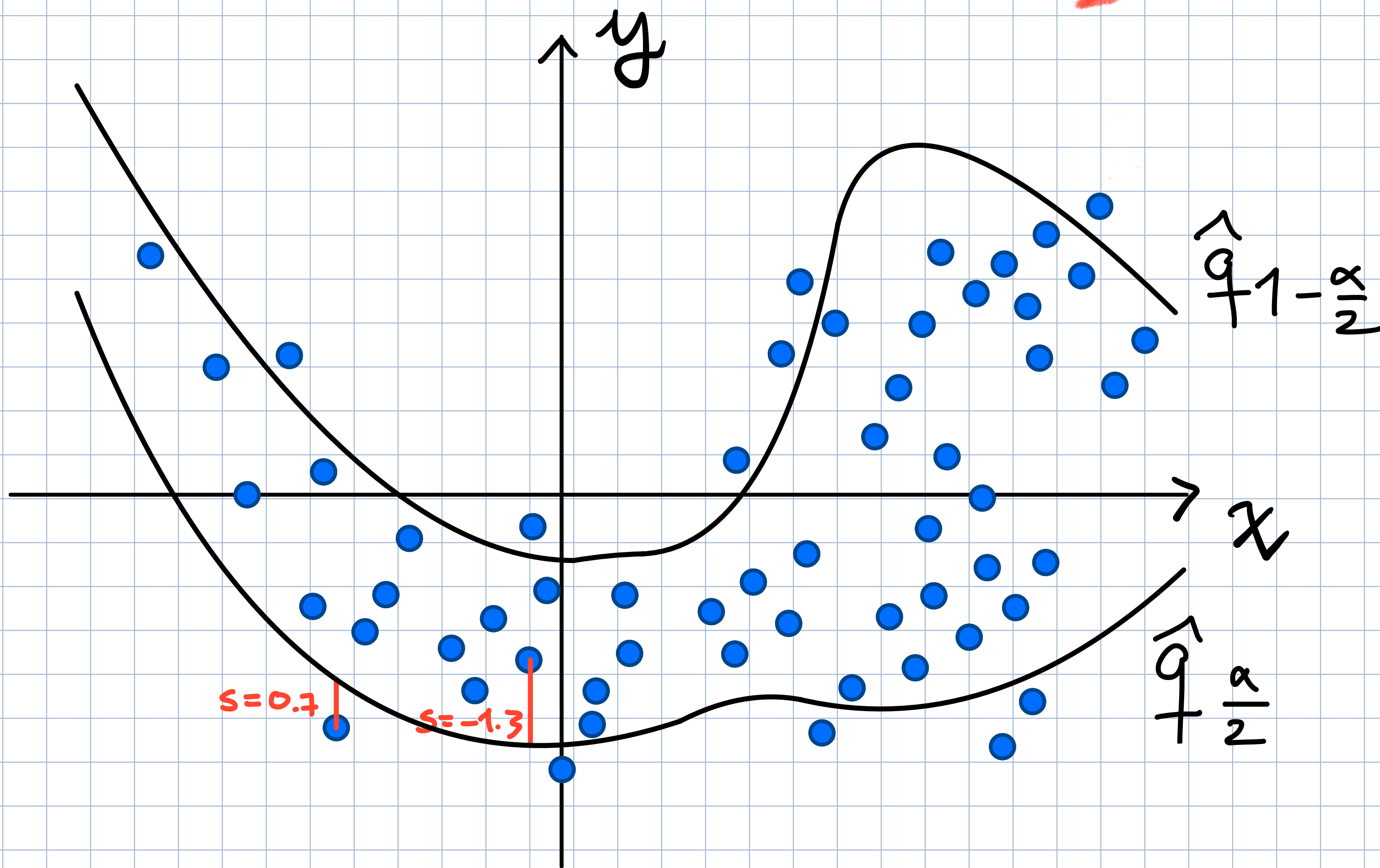
CQR

$$\mathbb{P}\left(Y \in \left[q_{\frac{\alpha}{2}}(X), q_{1-\frac{\alpha}{2}}(X)\right]\right) = ?$$



CQR: Calibration

$$s_i = \max\{ q_{\frac{\alpha}{2}}(x_i) - y_i ,\ y_i + q_{1-\frac{\alpha}{2}}(x_i) \}$$



CQR: Calibration

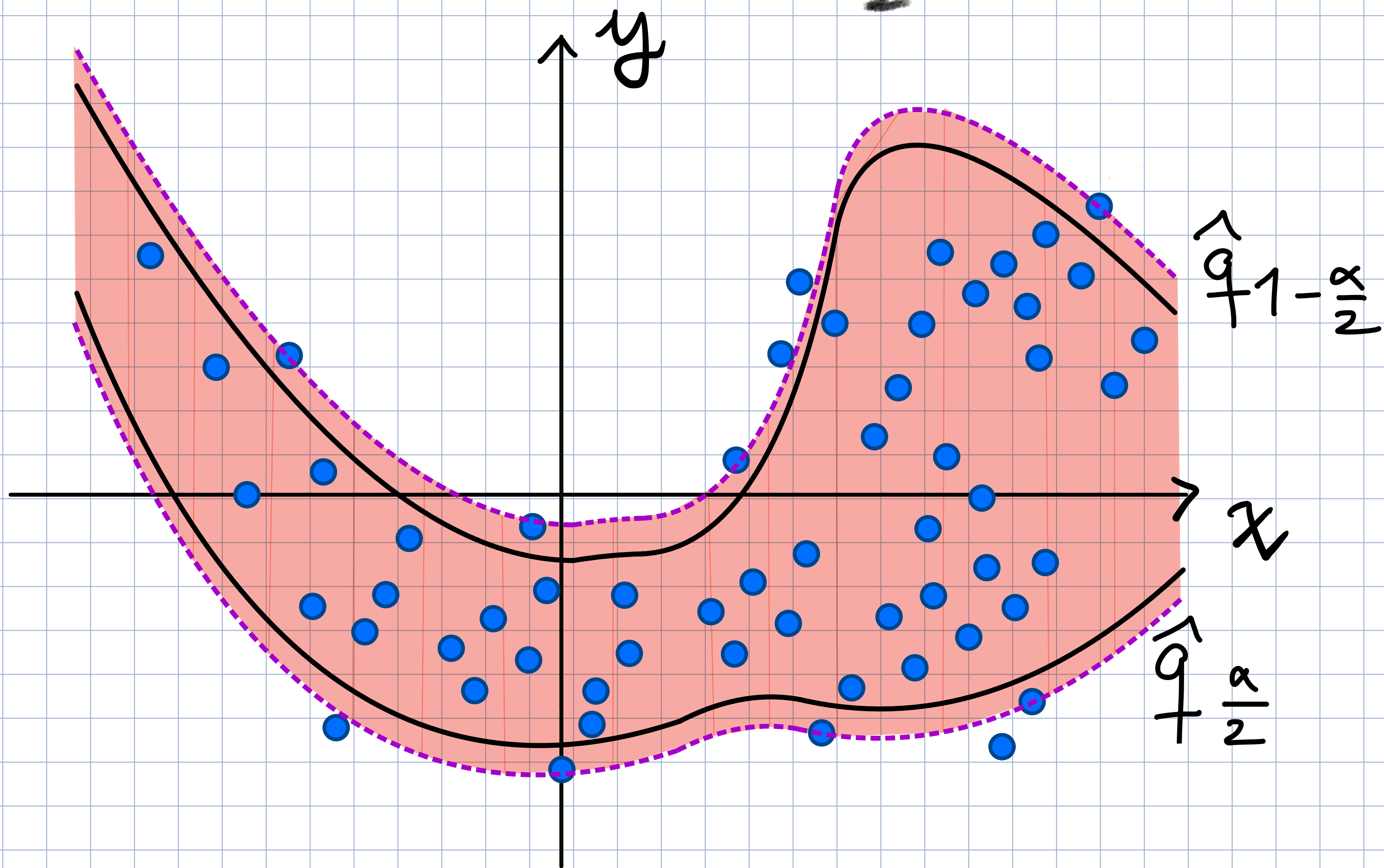
Calibration:

$$S_\alpha := \text{the } \left\lceil \frac{(n+1)(1-\alpha)}{n} \right\rceil\text{-th}$$

quantile of the scores $S_1, \ldots, S_n$

CQR : Calibration

$$C_\alpha(x) = \left[\hat{q}_{\frac{\alpha}{2}}(x) - \delta_\alpha, \hat{q}_{1-\frac{\alpha}{2}}(x) + \delta_\alpha\right]$$



Conformal Classification

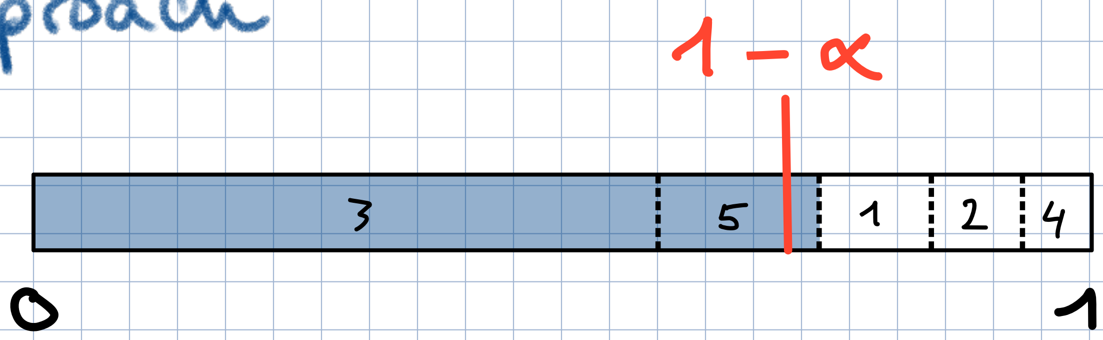
$$\hat{\pi}: \mathcal{X} \longrightarrow \mathcal{P}(\{1, \dots, K\})$$

$$x \longmapsto (\hat{\pi}_1(x), \dots, \hat{\pi}_K(x)\}$$

Rank: $\hat{\pi}_{(1)}(x) \geq \hat{\pi}_{(2)}(x) \geq \dots \geq \hat{\pi}_{(K)}(x)$

Naive approach

$\hat{\pi}(x)$



Conformal Classification

$$\hat{\pi} : \mathcal{X} \longrightarrow \mathcal{P}(\{1, \dots, K\})$$

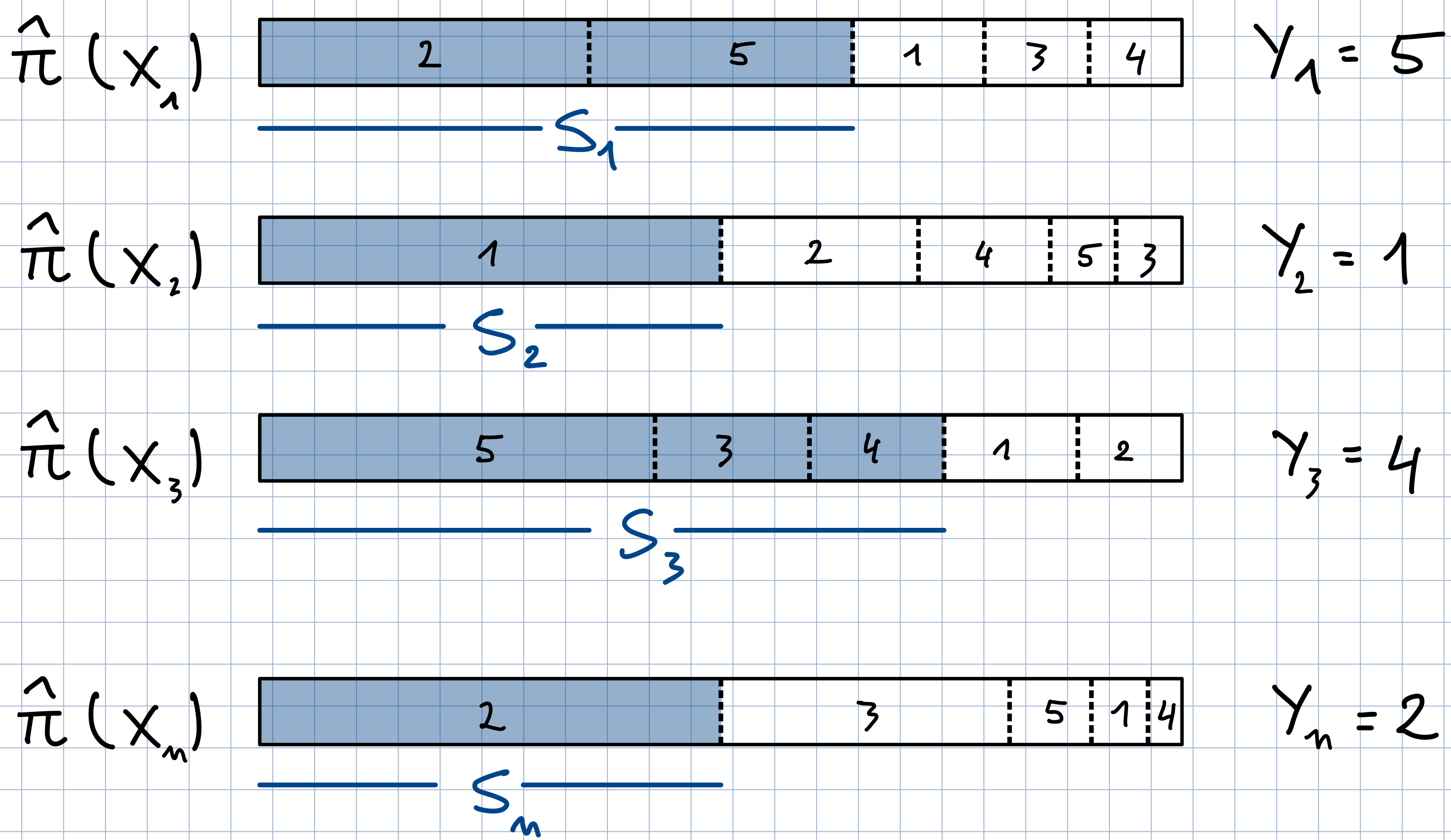
$$x \longmapsto (\hat{\pi}_1(x), \dots, \hat{\pi}_K(x)\}$$

Rank: $\hat{\pi}_{(1)}(x) \geq \hat{\pi}_{(2)}(x) \geq \dots \geq \hat{\pi}_{(K)}(x)$

$$L(x, \hat{\pi}, \tau) := \min_{c \in \{1, \dots, K\}} \left\{ \hat{\pi}_{(1)}(x) + \dots + \hat{\pi}_c(x) \geq \tau \right\}$$

Conformal Classification

Calibration:



CONFORMAL CLASSIFICATION

Calibration:

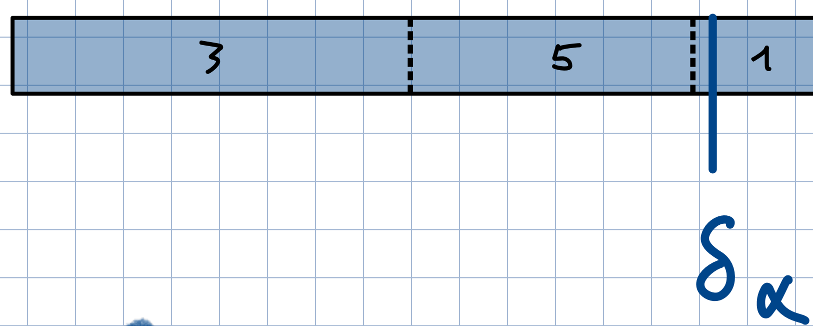
$$s_\alpha := \text{the } \left\lceil \frac{(n+1)(1-\alpha)}{n} \right\rceil\text{-th}$$

quantile of the scores $s_1, \ldots, s_n$

Conformal Classification

Inference:

$\hat{\pi}(x)$



$$\hat{C}_\alpha(x) = \{1, 3, 5\}$$

Guarantee:

$$\mathbb{P}\left(Y_{n+1} \in \hat{C}_\alpha(X_{n+1})\right) \geq 1 - \alpha$$