Mathematical Properties of Continuous Ranked Probability Score Forecasting

R. Pic[1], C. Dombry[1], P. Naveau[2] et M. Taillardat[3]

[1]Laboratoire de Mathématiques de Besançon, Université de Franche-Comté

[2]Laboratoire des Sciences du Climat et de l'Environnement, Université de Versailles Saint-Quentin

[3]Centre National de Recherches Météorologiques, Météo France

1. Probabilistic Forecasting
   - Context
   - Scoring Rules and Distributional Regression
   - CRPS
2. Statistical Learning
   - Theoretical Framework
   - Optimal Minimax Rate of Convergence
   - k-Nearest Neighbors
   - Kernel Method
3. Simulations

Table of Contents

1. Probabilistic Forecasting
   - Context
   - Scoring Rules and Distributional Regression
   - CRPS
2. Statistical Learning
   - Theoretical Framework
   - Optimal Minimax Rate of Convergence
   - k-Nearest Neighbors
   - Kernel Method
3. Simulations

Météo France

Probabilistic Forecasting

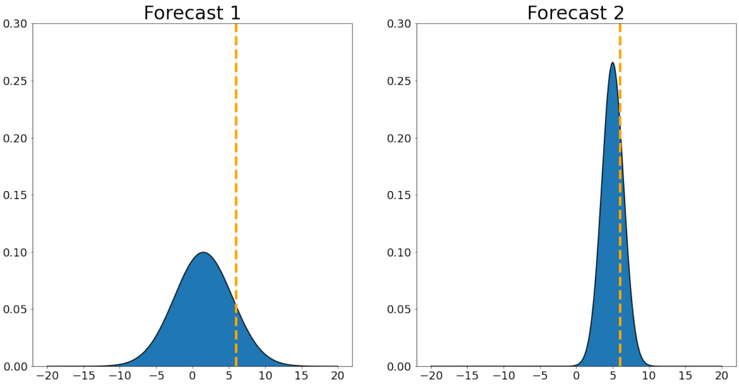
All those whose duty it is to issue regular daily forecasts know that there are times when they feel **very confident** and other times when they are **doubtful** as to coming weather. It seems to me that the condition of confidence or otherwise forms a **very important part of the prediction**.

W. Ernest Cook (MWR, 1906)

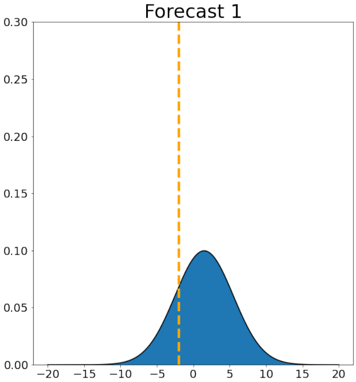
Comparing probabilistic forecasts and observation

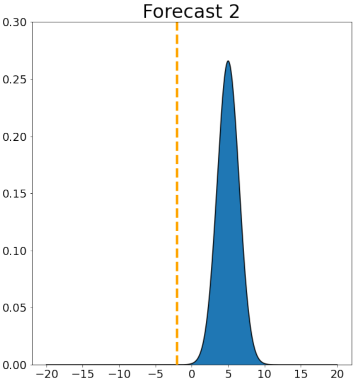
- How can we compare a distribution and an observation ?

- How can we compare a distribution and an observation ?



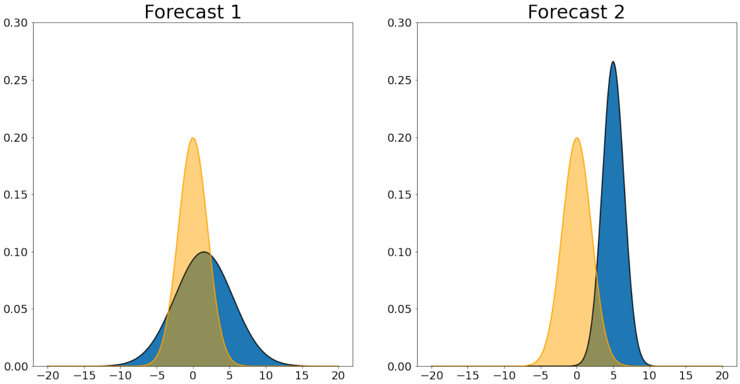
Comparing probabilistic forecasts and observation





- Which is the best forecast ?

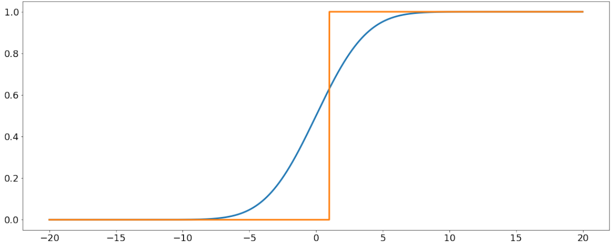
Comparing the predicted distribution and the real distribution



Continuous Ranked Probability Score

Continuous Ranked Probability Score (CRPS) : (Mateson, Wrinkler 1976)

$$\mathrm{CRPS}(F, y) = \int_{\mathbb{R}} (F(z) - \mathbb{1}_{y \leq z})^2 \mathrm{d}z$$



Difference of expected scores :

$$\begin{aligned}\overline{\mathrm{CRPS}}(F, G) - \overline{\mathrm{CRPS}}(G, G) &= \mathbb{E}_{Y \sim G}[\mathrm{CRPS}(F, Y) - \mathrm{CRPS}(G, Y)] \\ &= \int_{\mathbb{R}} (F(z) - G(z))^2 \mathrm{d}z\end{aligned}$$

$$\overline{\mathrm{CRPS}}(F, G) \geq \overline{\mathrm{CRPS}}(G, G)$$

Table of Contents

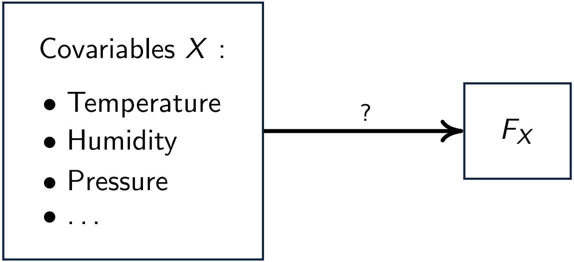
1 Probabilistic Forecasting
- Context
- Scoring Rules and Distributional Regression
- CRPS

2 Statistical Learning
- Theoretical Framework
- Optimal Minimax Rate of Convergence
- k-Nearest Neighbors
- Kernel Method

3 Simulations

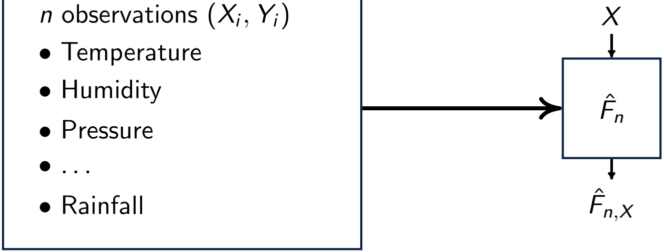
Theoretical framework

- $Y \in \mathbb{R}$ variable of interest, $X \in \mathbb{R}^d$ covariables with $(X, Y) \sim P$.
- Goal : estimate the conditional distribution of Y given X, noted $\mathbb{P}_{Y|X=x}(\mathrm{d}y)$.



Statistical Learning Framework

- In practice : estimate the conditional distribution of Y given X based on observations $D_n = \{(X_i, Y_i), i \in [\![1; n]\!]\}$ where (X_i, Y_i) are assumed i.i.d. following P.



- Evaluation via the CRPS : **expected risk**

$$R_P(\hat{F}_n) = \mathbb{E}_{D_n \sim P^n, (X,Y) \sim P}\left[\mathrm{CRPS}(\hat{F}_{n,X}, Y)\right]$$

- Rate of convergence for a given class of distributions ?
- Minimization of the maximal error on a class of distributions. (minimax error)

Definition

A sequence of positive numbers (a_n) is called an **optimal minimax rate of convergence** on the class $\mathcal{D}$ if

$$\liminf_{n\to\infty} \inf_{\hat{F}_n} \sup_{P\in\mathcal{D}} \frac{\mathbb{E}_{D_n\sim P^n}[R_P(\hat{F}_n)] - R_P(F^*)}{a_n} > 0 \tag{1}$$

and

$$\limsup_{n\to\infty} \inf_{\hat{F}_n} \sup_{P\in\mathcal{D}} \frac{\mathbb{E}_{D_n\sim P^n}[R_P(\hat{F}_n)] - R_P(F^*)}{a_n} < \infty, \tag{2}$$

where the infimum is taken over all distributional regression models $\hat{F}_n$ trained on D_n. If the sequence (a_n) satisfies only the lower bound (1), it is called a **lower minimax rate of convergence**.

Consider the following classes :

Definition

For $h \in (0,1]$, $C > 0$ and $M > 0$, let $\mathcal{D}^{(h,C,M)}$ be the class of distributions P such that $F_x^*(y) = P(Y \leq y | X = x)$ satisfies :

i) $X \in [0,1]^d$ P_X-a.s. ;

ii) For all $x \in [0,1]^d$, $\int_{\mathbb{R}} F_x^*(z)(1 - F_x^*(z))\mathrm{d}z \leq M$;

iii) $\|F_{x'}^* - F_x^*\|_{L^2} \leq C\|x' - x\|^h$ for all $x, x' \in [0,1]^d$.

Remark : Conditions similar to point regression (Györfi et al., 2002).

Class of distributions

i) $\boldsymbol{X \in [0,1]^d}$ $\boldsymbol{P_X}$**-a.s. ;**
$\rightarrow$ More generally a compact.

i) $\boldsymbol{X \in [0,1]^d}$ $\boldsymbol{P_X}$**-a.s. ;**
$\rightarrow$ More generally a compact.

ii) For all $\boldsymbol{x \in [0,1]^d}$**,** $\boldsymbol{\int_{\mathbb{R}} F_x^*(z)(1 - F_x^*(z))\mathrm{d}z \leq M}$ **;**
$\rightarrow$ The dispersion of $Y|X = x$ remains bounded for all $x \in [0,1]^d$.

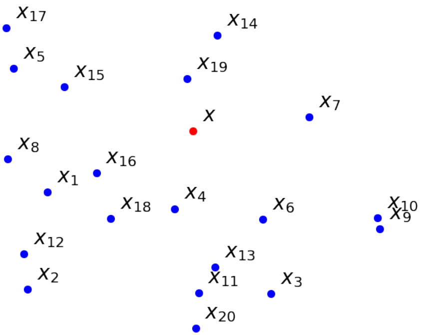
i) $\boldsymbol{X \in [0,1]^d}$ $\boldsymbol{P_X}$**-a.s. ;**
$\rightarrow$ More generally a compact.

ii) **For all** $\boldsymbol{x \in [0,1]^d}$, $\boldsymbol{\int_{\mathbb{R}} F_x^*(z)(1 - F_x^*(z))\mathrm{d}z \leq M}$ **;**
$\rightarrow$ The dispersion of $Y|X = x$ remains bounded for all $x \in [0,1]^d$.

iii) $\boldsymbol{\|F_{x'}^* - F_x^*\|_{L^2} \leq C\|x' - x\|^h}$ **for all** $\boldsymbol{x, x' \in [0,1]^d}$.

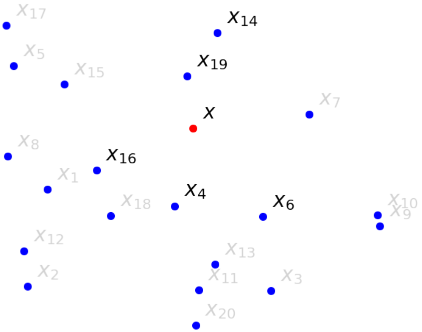
$$
\begin{aligned}
\mathbb{E}_{D_n \sim P^n}[R_P(\hat{F}_n)] - R_P(F^*) &= \mathbb{E}_{D_n \sim P^n, (X,Y) \sim P}[\mathrm{CRPS}(\hat{F}_{n,X}, Y) - \mathrm{CRPS}(F_X^*, Y)] \\
&= \mathbb{E}_{D_n \sim P^n, X \sim P_X}\left[\int_{\mathbb{R}} |\hat{F}_{n,X}(z) - F_X^*(z)|^2 \mathrm{d}z\right] \\
&= \mathbb{E}_{D_n \sim P^n, X \sim P_X}\left[\|\hat{F}_{n,X} - F_X^*\|_{L^2}\right]
\end{aligned}
$$

Using knowledge from previous observations at $X = x_i$ to extrapolate the value at $X = x$ $\rightarrow$ Need regularity of F^*.



$$\hat{F}_{n,x}(z) = \frac{1}{k_n}\sum_{i=1}^{k_n} \mathbb{1}_{y_{i:n}(x)\leq z}, \quad {}_{i:n}(x) \text{ index of the } i\text{-th nearest neighbor of } x.$$

k-Nearest Neighbors



$$\hat{F}_{n,x}(z) = \frac{1}{k_n} \sum_{i=1}^{k_n} \mathbb{1}_{y_{i:n}(x) \leq z}, \quad {}_{i:n}(x) \text{ index of the } i\text{-th nearest neighbor of } x.$$

Proposition

Assume $P \in \mathcal{D}^{(h,C,M)}$ *and let* $\hat{F}_n$ *be the k-NN model. Then,*

$$\mathbb{E}_{D_n \sim P^n}[R_P(\hat{F}_n)] - R_P(F^*) \leq \begin{cases} 8^h C^2 \left(\frac{k_n}{n}\right)^h + \frac{M}{k_n} & \text{if } d = 1, \\ {c_d}^h C^2 \left(\frac{k_n}{n}\right)^{2h/d} + \frac{M}{k_n} & \text{if } d \geq 2, \end{cases}$$

where $c_d = \frac{2^{3+\frac{2}{d}}(1+\sqrt{d})^2}{V_d^{2/d}}$ *and* V_d *is the volume of the unit ball in* $\mathbb{R}^d$.

Theorem

For $d \geq 2$, the optimal minimax rate of convergence on the class $\mathcal{D}^{(h,C,M)}$ is $\boldsymbol{a_n = n^{-\frac{2h}{2h+d}}}$. Moreover, the k-NN algorithm reaches the optimal rate of convergence for

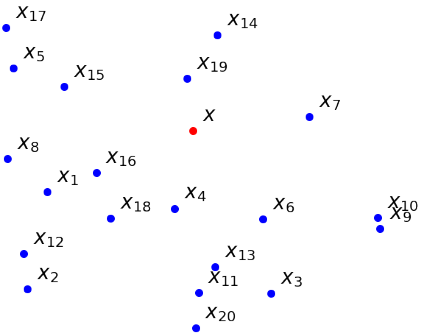
$$k_n = \left(\frac{Md}{2hC^2 c_d^h} \right)^{\frac{d}{2h+d}} n^{\frac{2h}{2h+d}}.$$

Theorem

For $d \geq 2$, the optimal minimax rate of convergence on the class $\mathcal{D}^{(h,C,M)}$ is $\boldsymbol{a_n = n^{-\frac{2h}{2h+d}}}$. Moreover, the k-NN algorithm reaches the optimal rate of convergence for

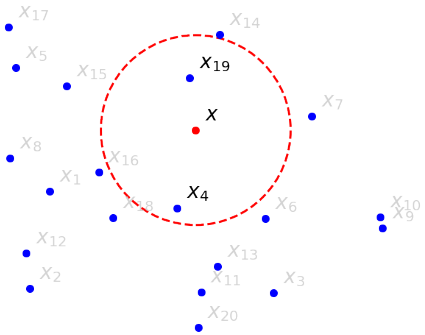
$$k_n = \left(\frac{Md}{2hC^2c_d^h}\right)^{\frac{d}{2h+d}} n^{\frac{2h}{2h+d}}.$$

- What happens in $d = 1$?
- Interesting result but k-NN not used in practice.



$$\hat{F}_{n,x}(z) = \frac{\sum_{i=1}^{n} K(\frac{x-x_i}{h_n})\mathbb{1}_{y_i \le z}}{\sum_{i=1}^{n} K(\frac{x-x_i}{h_n})}, \text{ with } K(z) = \mathbb{1}_{\{\|z\| \le 1\}}.$$

Kernel Method



$$\hat{F}_{n,x}(z) = \frac{\sum_{i=1}^{n} K(\frac{x-x_i}{h_n})\mathbb{1}_{y_i \leq z}}{\sum_{i=1}^{n} K(\frac{x-x_i}{h_n})}, \text{ with } K(z) = \mathbb{1}_{\{\|z\| \leq 1\}}.$$

Proposition

Assume $P \in \mathcal{D}^{(h,C,M)}$ and let $\hat{F}_n$ be the naive kernel model. Then,

$$\mathbb{E}_{D_n \sim P^n}[R_P(\hat{F}_n)] - R_P(F^*) \leq \tilde{c}_d \frac{2M + Cd^{h/2} + \frac{M}{n}}{nh_n^d} + C^2 h_n^{2h}$$

where $\tilde{c}_d$ only depends on $d \geq 1$.

Theorem

For any d, the optimal minimax rate of convergence on the class $\mathcal{D}^{(h,C,M)}$ is $\boldsymbol{a_n = n^{-\frac{2h}{2h+d}}}$. Moreover, the naive kernel algorithm reaches the optimal rate of convergence for $h_n = \left(\frac{\tilde{c}_d d(M + Cd^{h/2} + \frac{M}{n})}{2hC^2} \right)^{\frac{1}{2h+d}} n^{-\frac{1}{2h+d}}$.

Proposition

Assume $P \in \mathcal{D}^{(h,C,M)}$ and let $\hat{F}_n$ be the naive kernel model. Then,

$$\mathbb{E}_{D_n \sim P^n}[R_P(\hat{F}_n)] - R_P(F^*) \leq \tilde{c}_d \frac{2M + Cd^{h/2} + \frac{M}{n}}{nh_n^d} + C^2 h_n^{2h}$$

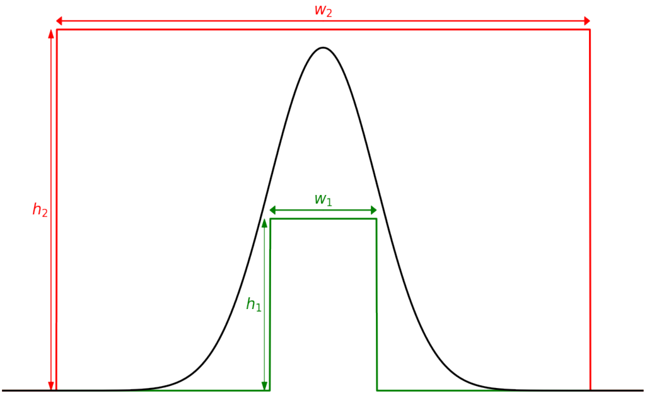
where $\tilde{c}_d$ only depends on $d \geq 1$.

Theorem

For any d, the optimal minimax rate of convergence on the class $\mathcal{D}^{(h,C,M)}$ is $\boldsymbol{a_n = n^{-\frac{2h}{2h+d}}}$. Moreover, the naive kernel algorithm reaches the optimal rate of convergence for $h_n = \left(\frac{\tilde{c}_d d(M + Cd^{h/2} + \frac{M}{n})}{2hC^2}\right)^{\frac{1}{2h+d}} n^{-\frac{1}{2h+d}}$.

- Takes care of the $d = 1$ case.
- More used in practice ?

Boxed Kernel



1 Probabilistic Forecasting
- Context
- Scoring Rules and Distributional Regression
- CRPS

2 Statistical Learning
- Theoretical Framework
- Optimal Minimax Rate of Convergence
- k-Nearest Neighbors
- Kernel Method

3 Simulations

Simulations

- $X_1, X_2 \sim \mathcal{U}([0,1])$
- $Y = X_1 + X_2 + \sigma\epsilon$ with $\epsilon \sim \mathcal{N}(0,1)$
- $Y|X \sim \mathcal{N}(X_1 + X_2, \sigma^2)$

Simulations

- $X_1, X_2 \sim \mathcal{U}([0,1])$
- $Y = X_1 + X_2 + \sigma\epsilon$ with $\epsilon \sim \mathcal{N}(0,1)$
- $Y|X \sim \mathcal{N}(X_1 + X_2, \sigma^2)$
- Checking the conditions :

- $X_1, X_2 \sim \mathcal{U}([0,1])$
- $Y = X_1 + X_2 + \sigma\epsilon$ with $\epsilon \sim \mathcal{N}(0,1)$
- $Y|X \sim \mathcal{N}(X_1 + X_2, \sigma^2)$
- Checking the conditions :
  i) $\boldsymbol{X \in [0,1]^d}$ **$\boldsymbol{P_X}$-a.s. ;** ✓
  ii) **For all** $\boldsymbol{x \in [0,1]^d}$**,** $\boldsymbol{\int_{\mathbb{R}} F_x^*(z)(1 - F_x^*(z))\mathrm{d}z \leq M}$ **;**
  $\rightarrow M = \frac{\sigma}{\sqrt{\pi}}$ ✓
  iii) $\boldsymbol{\|F_{x'}^* - F_x^*\|_{L^2} \leq C\|x' - x\|^h}$ **for all** $\boldsymbol{x, x' \in [0,1]^d}$.
  $\rightarrow$ Hard to get optimal values for C and h but $h = 1$ works. ✓

- $X_1, X_2 \sim \mathcal{U}([0,1])$
- $Y = X_1 + X_2 + \sigma\epsilon$ with $\epsilon \sim \mathcal{N}(0,1)$
- $Y|X \sim \mathcal{N}(X_1 + X_2, \sigma^2)$
- Checking the conditions :
  i) $\boldsymbol{X \in [0,1]^d}$ **$\boldsymbol{P_X}$-a.s. ;** ✓
  ii) **For all** $\boldsymbol{x \in [0,1]^d}$, $\boldsymbol{\int_{\mathbb{R}} F_x^*(z)(1 - F_x^*(z))\mathrm{d}z \leq M}$ **;**
  $\rightarrow M = \frac{\sigma}{\sqrt{\pi}}$ ✓
  iii) $\boldsymbol{\|F_{x'}^* - F_x^*\|_{L^2} \leq C\|x' - x\|^h}$ **for all** $\boldsymbol{x, x' \in [0,1]^d}$.
  $\rightarrow$ Hard to get optimal values for C and h but $h = 1$ works. ✓
- k-NN :

$$\hat{F}_{n,x}(z) = \frac{1}{k_n} \sum_{i=1}^{k_n} \mathbb{1}_{y_{i:n}(x) \leq z}$$

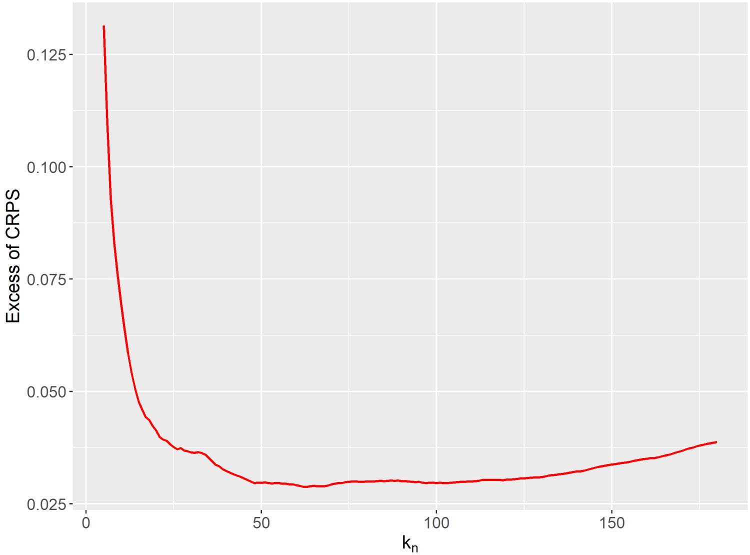
- $X_1, X_2 \sim \mathcal{U}([0,1])$
- $Y = X_1 + X_2 + \sigma\epsilon$ with $\epsilon \sim \mathcal{N}(0,1)$
- $Y|X \sim \mathcal{N}(X_1 + X_2, \sigma^2)$
- Checking the conditions :
  - **i)** $\boldsymbol{X \in [0,1]^d}$ **$\boldsymbol{P_X}$-a.s. ;** ✓
  - **ii)** **For all** $\boldsymbol{x \in [0,1]^d}$**,** $\boldsymbol{\int_{\mathbb{R}} F_x^*(z)(1 - F_x^*(z))\mathrm{d}z \leq M}$ **;**
    $\rightarrow M = \frac{\sigma}{\sqrt{\pi}}$ ✓
  - **iii)** $\boldsymbol{\|F_{x'}^* - F_x^*\|_{L^2} \leq C\|x' - x\|^h}$ **for all** $\boldsymbol{x, x' \in [0,1]^d}$.
    $\rightarrow$ Hard to get optimal values for C and h but $h = 1$ works. ✓
- k-NN :

$$\hat{F}_{n,x}(z) = \frac{1}{k_n} \sum_{i=1}^{k_n} \mathbb{1}_{y_{i:n}(x) \leq z}$$

$$\overline{\mathrm{CRPS}}(F_{n,x}, F_x^*) - \overline{\mathrm{CRPS}}(F_x^*, F_x^*) = \int_{\mathbb{R}} \left(\frac{1}{k_n} \sum_{i=1}^{k_n} \mathbb{1}_{y_{i:n}(x) \leq z} - \Phi\left(\frac{z - (x_1 + x_2)}{\sigma}\right)\right)^2 \mathrm{d}z$$

CRPS vs. k_n

Parameters : $\sigma = 1$ and $n = 200$.

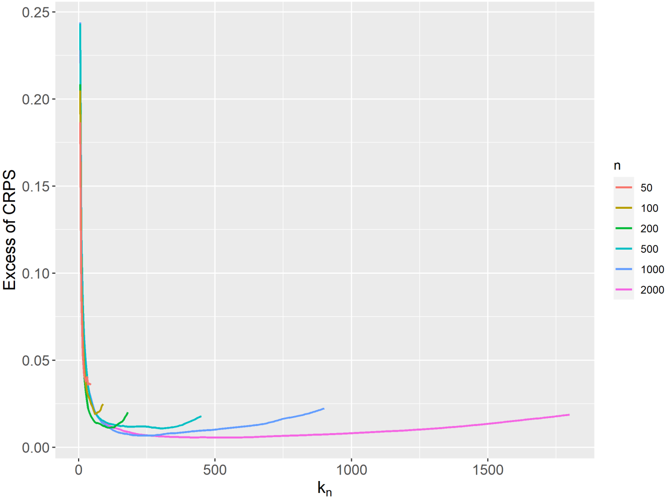


Scaling of k_n with n, $\sigma = 2$

$$k_n \propto n^{\frac{2h}{2h+d}}$$

Scaling of k_n with n, $\sigma = 2$

$$k_n \propto n^{\frac{2h}{2h+d}}$$



Scaling of k_n with n, $\sigma = 2$

$$k_n \propto n^{\frac{2h}{2h+d}}$$

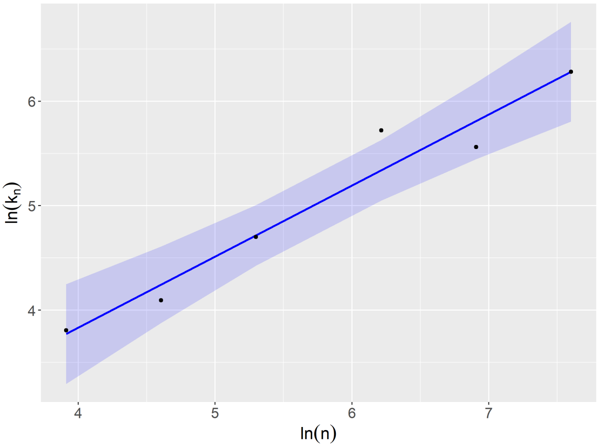


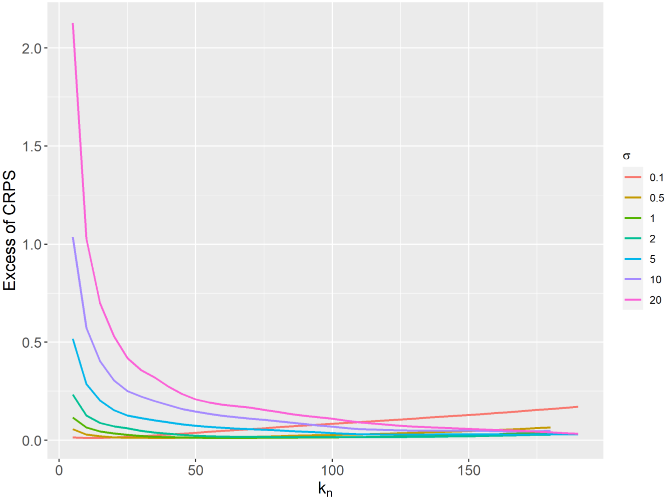
Figure – Equation : $y = 1.1 + 0.68x,\ R^2 = 0.952$

Scaling of k_n with σ, $n = 200$

$$k_n \propto M^{\frac{d}{2h+d}} \propto \sigma^{\frac{d}{2h+d}}$$

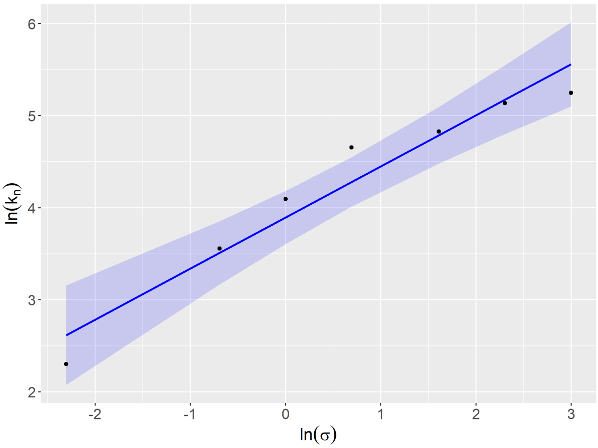
Scaling of k_n with σ, $n = 200$

$$k_n \propto M^{\frac{d}{2h+d}} \propto \sigma^{\frac{d}{2h+d}}$$



Scaling of k_n with σ, $n = 200$

$$k_n \propto M^{\frac{d}{2h+d}} \propto \sigma^{\frac{d}{2h+d}}$$



Conclusion

- Optimal minimax rate of convergence for distributional regression.
- Upper bound on the convergence rate for k-NN and kernel methods at fixed n.
- Perspectives :
    - Study other algorithms : Random Forests (e.g. QRF).
    - Study other definitions of convergence : other distances.

Preprint : Mathematical Properties of Continuous Ranked Probability Score Forecasting, Pic et al. (`https://arxiv.org/abs/2205.04360`)

Upper Bound

$$\mathbb{E}[R_P(\hat{F}_n)] - R_P(F^*) = \mathbb{E}\left[\int_{\mathbb{R}} |\hat{F}_{n,x}(z) - F_x^*(z)|^2 \mathrm{d}z\right]$$

Upper Bound

$$\mathbb{E}[R_P(\hat{F}_n)] - R_P(F^*) = \mathbb{E}\left[\int_{\mathbb{R}} |\hat{F}_{n,x}(z) - F_x^*(z)|^2 \mathrm{d}z\right]$$

$$\begin{aligned}
&\mathbb{E}\big[|\hat{F}_{n,x}(z) - F_x^*(z)|^2\big] \\
&= \underbrace{\mathbb{E}\Big[\Big(\frac{1}{k_n}\sum_{i=1}^{k_n}\big(F^*_{X_{i:n}(x)}(z) - F_x^*(z)\big)\Big)^2\Big]}_{\text{squared bias}} + \underbrace{\frac{1}{k_n^2}\sum_{i=1}^{k_n}\mathbb{E}\Big[F^*_{X_{i:n}(x)}(z)(1 - F^*_{X_{i:n}(x)}(z))\Big]}_{\text{variance}}
\end{aligned}$$

Upper Bound

$$\mathbb{E}[R_P(\hat{F}_n)] - R_P(F^*) = \mathbb{E}\left[\int_{\mathbb{R}} |\hat{F}_{n,x}(z) - F^*_x(z)|^2 \mathrm{d}z\right]$$

$$\mathbb{E}\left[|\hat{F}_{n,x}(z) - F^*_x(z)|^2\right]$$
$$= \underbrace{\mathbb{E}\Big[\Big(\frac{1}{k_n}\sum_{i=1}^{k_n}\big(F^*_{X_{i:n(x)}}(z) - F^*_x(z)\big)\Big)^2\Big]}_{\text{squared bias}} + \underbrace{\frac{1}{k_n^2}\sum_{i=1}^{k_n}\mathbb{E}\Big[F^*_{X_{i:n(x)}}(z)(1 - F^*_{X_{i:n(x)}}(z))\Big]}_{\text{variance}}$$

Integrating and using Jensen's inequality :

$$\mathbb{E}[R_P(\hat{F}_n)] - R_P(F^*)$$
$$\leq \frac{1}{k_n}\sum_{i=1}^{k_n}\mathbb{E}\left[\int_{\mathbb{R}}(F^*_{X_{i:n}(X)}(z) - F^*_X(z))^2\mathrm{d}z\right] + \frac{1}{k_n^2}\sum_{i=1}^{k_n}\mathbb{E}\left[\int_{\mathbb{R}} F^*_{X_{i:n}(X)}(z)(1 - F^*_{X_{i:n}(X)}(z))\mathrm{d}z\right].$$

Upper Bound

$$\mathbb{E}[R_P(\hat{F}_n)] - R_P(F^*) = \mathbb{E}\left[\int_{\mathbb{R}} |\hat{F}_{n,x}(z) - F^*_x(z)|^2 \mathrm{d}z\right]$$

$$\mathbb{E}\big[|\hat{F}_{n,x}(z) - F^*_x(z)|^2\big]$$
$$= \underbrace{\mathbb{E}\Big[\Big(\frac{1}{k_n}\sum_{i=1}^{k_n}\big(F^*_{X_{i:n}(x)}(z) - F^*_x(z)\big)\Big)^2\Big]}_{\text{squared bias}} + \underbrace{\frac{1}{k_n^2}\sum_{i=1}^{k_n}\mathbb{E}\Big[F^*_{X_{i:n}(x)}(z)(1 - F^*_{X_{i:n}(x)}(z))\Big]}_{\text{variance}}$$

Integrating and using Jensen's inequality :

$$\mathbb{E}[R_P(\hat{F}_n)] - R_P(F^*)$$
$$\leq \frac{1}{k_n}\sum_{i=1}^{k_n}\mathbb{E}\left[\int_{\mathbb{R}}(F^*_{X_{i:n}(X)}(z) - F^*_X(z))^2\mathrm{d}z\right] + \frac{1}{k_n^2}\sum_{i=1}^{k_n}\mathbb{E}\left[\int_{\mathbb{R}} F^*_{X_{i:n}(X)}(z)(1 - F^*_{X_{i:n}(X)}(z))\mathrm{d}z\right].$$

Using conditions *ii*) and *iii*) :

$$\mathbb{E}[R_P(\hat{F}_n)] - R_P(F^*) \leq C^2\mathbb{E}\big[\|X_{k_n:n}(X) - X\|^{2h}\big] + \frac{M}{k_n}$$

Upper Bound

Biau & Devroye (2015) :

$$\mathbb{E}[\|X_{k_n:n}(X) - X\|^2] \leq \begin{cases} 8\dfrac{k_n}{n} & \text{if } d = 1, \\ c_d \left(\dfrac{k_n}{n}\right)^{2/d} & \text{if } d \geq 2. \end{cases}$$

Upper Bound

Biau & Devroye (2015) :

$$\mathbb{E}[\|X_{k_n:n}(X) - X\|^2] \leq \begin{cases} 8\dfrac{k_n}{n} & \text{if } d = 1, \\ c_d \left(\dfrac{k_n}{n}\right)^{2/d} & \text{if } d \geq 2. \end{cases}$$

Finally,

$$\mathbb{E}[R_P(\hat{F}_n)] - R_P(F^*) \leq \begin{cases} C^2 8^h \left(\dfrac{k_n}{n}\right)^h + \dfrac{M}{k_n} & \text{if } d = 1, \\ C^2 {c_d}^h \left(\dfrac{k_n}{n}\right)^{2h/d} + \dfrac{M}{k_n} & \text{if } d \geq 2, \end{cases}$$