Sequential Quantile Prediction of Time Series.
Joined work with Gérard Biau.

Benoît Patra, benoit.patra@lokad.com

JDS BORDEAUX, MAY 2009.
## Time series prediction.
- **Time series** prediction has a long history (Yule, 1927).
- **Parametric approaches** (Until 70’s).
- Recently **non parametric** approach.

## Quantile forecasting.
Given a stochastic process $Y_1, Y_2, \ldots$:
- Usually, estimate the **conditional mean** of $Y_n$ given $Y_1, \ldots, Y_{n-1}$.
- Here: the **conditional $\tau$th quantile** of $Y_n$ given $Y_1, \ldots, Y_{n-1}$.
Time series prediction.

- Time series prediction has a long history \((Yule, 1927)\).
- Parametric approaches (Until 70’s).
- Recently non parametric approach.

Quantile forecasting.

Given a stochastic process \(Y_1, Y_2, \ldots\).

- Usually, estimate the conditional mean of \(Y_n\) given \(Y_1, \ldots, Y_{n-1}\).
- Here: the conditional \(\tau\)th quantile of \(Y_n\) given \(Y_1, \ldots, Y_{n-1}\).
Introduction.

What for?
- Understand conditional distributions.
- $\tau = 0.5$ robust forecasting.
- Build confidence interval.

Applications fields.
- Finance: CVAR. Also biology, medicine, telecoms...
- Here: call volumes (optimize staff in a call center).
Introduction.

What for?
- Understand conditional distributions.
- $\tau = 0.5$ robust forecasting.
- Build confidence interval.

Applications fields.
- Finance: CVAR. Also biology, medicine, telecoms...
- Here: call volumes (optimize staff in a call center).
Figure: Quantile forecast with $\tau = 0.1, 0.9$. 
Quantile Regression.

Conditional quantiles.

\( X \) multivariate random variable, \( Y \) real valued random variable,

\[ q_{\tau}(X) \triangleq F_{Y|X}^{-1}(\tau) = \inf\{ t \in \mathbb{R} : F_{Y|X}(t) \geq \tau \}. \]

\( F_{Y|X} \) conditional cumulative distribution function.

Proposition (Koenker, 2005)

\[ q_{\tau}(X) \in \arg\min_{q(.) \in \mathbb{R}} \mathbb{E}_{P_{Y|X}} [\rho_{\tau}(Y - q(X))]. \]
Quantile Regression.

Conditional quantiles.

$X$ multivariate random variable, $Y$ real valued random variable,

$$q_\tau(X) \triangleq F_{Y|X}^{-1}(\tau) = \inf\{t \in \mathbb{R} : F_{Y|X}(t) \geq \tau\}.$$  

$F_{Y|X}$ conditional cumulative distribution function.

**Proposition (Koenker, 2005)**

$$q_\tau(X) \in \arg\min_{q(.) \in \mathbb{R}} \mathbb{E}_{P_{Y|X}} [\rho_\tau (Y - q(X))].$$
Figure: Pinball function graph.
On line.

- Consider the sequential (= on-line) quantile prediction of time series.
- Including series that do not necessarily satisfy classical statistical assumptions for bounded, mixing or Markovian process.

Goal.

- Show consistency results under a minimum of hypotheses.

Notation.

- \( y_1^n = (y_1, \ldots, y_n) \) real sequence.
- \( Y_1^n = (Y_1, \ldots, Y_n) \) random variables sequence.
On line.

- Consider the sequential (= on-line) quantile prediction of time series.
- Including series that do not necessarily satisfy classical statistical assumptions for bounded, mixing or Markovian process.

Goal.

- Show consistency results under a minimum of hypotheses.

Notation.

- $y_1^n = (y_1, \ldots, y_n)$ real sequence.
- $Y_1^n = (Y_1, \ldots, Y_n)$ random variables sequence.
Non parametric framework.

On line.

- Consider the sequential (= on-line) quantile prediction of time series.
- Including series that do not necessarily satisfy classical statistical assumptions for bounded, mixing or Markovian process.

Goal.

- Show consistency results under a minimum of hypotheses.

Notation.

- \( y_1^n = (y_1, \ldots, y_n) \) real sequence.
- \( Y_1^n = (Y_1, \ldots, Y_n) \) random variables sequence.
Non parametric framework.

On line.

- Consider the sequential (= on-line) quantile prediction of time series.
- Including series that do not necessarily satisfy classical statistical assumptions for bounded, mixing or Markovian process.

Goal.

- Show consistency results under a minimum of hypotheses.

Notation.

- \( y_1^n = (y_1, \ldots, y_n) \) real sequence.
- \( Y_1^n = (Y_1, \ldots, Y_n) \) random variables sequence.
Non parametric framework.

Framework.
- Here, we observe a string realization $y_1^{n-1}$ of a stationary and ergodic process $\{Y_n\}_{-\infty}^{\infty} \ldots$
- ... and try to estimate $q_\tau(y_1^{n-1}) = F_{Y_n | Y_1^{n-1}=y_1^{n-1}}(\tau)$, the conditional quantile at time $n$.

Strategy.
Sequence $g = \{g_n\}_{n=1}^\infty$ of $\tau$th quantile forecasting functions $g_n : \mathbb{R}^{n-1} \rightarrow \mathbb{R}$.
Quantile estimation at time $n$ is $g_n(y_1^{n-1})$. 
Here, we observe a string realization $y_{1}^{n-1}$ of a stationary and ergodic process $\{Y_{n}\}_{-\infty}^{\infty}$.

... and try to estimate $q_{\tau}(y_{1}^{n-1}) = F_{Y_{n}|Y_{1}^{n-1}=y_{1}^{n-1}(\tau)}$, the conditional quantile at time $n$.

**Strategy.**

Sequence $g = \{g_{n}\}_{n=1}^{\infty}$ of $\tau$th quantile forecasting functions

$$g_{n} : \mathbb{R}^{n-1} \rightarrow \mathbb{R}.$$
At time \( n \) the cumulative pinball error of the strategy \( g \) is

\[
L_n(g) = \frac{1}{n} \sum_{t=1}^{n} \rho_{\tau} \left( y_t - g_t(y_{t-1}) \right).
\]

A fundamental limit (Algoët, 1994).

For any stationary and ergodic process \( \{Y_n\}_{n=\infty}^{+\infty} \),

\[
\liminf_{n \to \infty} L_n(g) \geq L^* \quad \text{a.s.,}
\]

where

\[
L^* = \mathbb{E} \min_{q(\cdot)} \mathbb{E}_{\mathbb{P}}_{Y_0|Y_{-\infty}^{-1}} \left[ \rho_{\tau} \left( Y_0 - q(Y_{-\infty}^{-1}) \right) \right].
\]
Errors.

Empirical measure criterion.

At time $n$ the **cumulative pinball error** of the strategy $g$ is

$$L_n(g) = \frac{1}{n} \sum_{t=1}^{n} \rho_\tau \left( y_t - g_t(y_1^{t-1}) \right).$$

A fundamental limit (**Algoët, 1994**).

For any stationary and ergodic process $\{ Y_n \}_{n=-\infty}^{+\infty}$,

$$\lim_{n \to \infty} \inf L_n(g) \geq L^* \quad \text{a.s.},$$

where

$$L^* = \mathbb{E} \left[ \min_{q(.)} \mathbb{E}_P \mathbb{E}_{Y_0 \mid Y_{-\infty}^{1}} \left[ \rho_\tau \left( Y_0 - q(Y_{-\infty}^{1}) \right) \right] \right].$$
A NN based aggregation scheme.

On line learning.

Scheme inspired from prediction of individual sequences.


Previous works.

- ... prediction of time series (fourth moment). Biau, Beakley, Györfi, Ottucsàk, 2009.
A NN based aggregation scheme.

On line learning.
Scheme inspired from prediction of individual sequences.

Previous works.
- ... prediction of time series (fourth moment). Biau, Beakley, Györfi, Ottucsàk, 2009.
A NN based aggregation scheme.

On line learning.

Scheme inspired from prediction of individual sequences.

Previous works.

- ... prediction of time series (fourth moment). Biau, Beakley, Györfi, Ottucsàk, 2009.
Nearest neighbors strategy.

**Elementary predictors.**
- Define infinite array of experts $h_n^{(k,\ell)}$: $k, \ell = 1, 2, \ldots$

**Each expert has a job.**
- At time $n$, expert $h_n^{(k,\ell)}$ searches for the $\ell$ nearest neighbors of length $k$. 
Nearest neighbors strategy.

Elementary predictors.
- Define infinite array of experts $h_{n}^{(k,\ell)}$: $k, \ell = 1, 2, \ldots$

Each expert has a job.
- At time $n$, expert $h_{n}^{(k,\ell)}$ searches for the $\bar{\ell}$ nearest neighbors of length $k$. 
| 1,34 | 1,78 | 2,56 | 1,88 | 0,57 | −1,25 | 0,19 | 3,18 | 4,13 | 2,22 | 1,34 | 0,26 | −1,90 | −2,29 | 0,88 | 1,28 | 3,31 | 4,12 | 5,15 | 3,31 |
| 2,27 | 2,89 | 2,12 | 1,78 | 2,67 | −3,16 | 0,01 | 1,16 | 5,17 | 6,17 | 7,18 | 9,10 | 8,18 | 1,16 | 5,17 | 6,17 | 5,15 | 3,14 | 2,18 | 1,18 | 0,99 |
| 0,10 | 1,15 | 2,17 | 3,72 | −1,71 | 6,39 | 5,16 | 3,13 | 1,89 | 0,90 | 0,91 | 0,11 | −0,20 | 1,89 | 2,84 | 3,92 | 2,99 | 2,21 | 1,73 |

**Figure:** Work of fundamental expert with $k = 3$ and $\ell = 4$. 
| 1.34 | 1.78 | 2.56 | 1.88 | 0.57 | -1.25 | 0.19 | 3.18 | 4.13 | 2.22 | 1.34 | 0.26 | -1.90 | -2.29 | 0.88 | 1.28 | 3.31 | 4.12 | 5.15 | 3.31 |
| 2.27 | 2.89 | 2.12 | 1.78 | 2.67 | -3.16 | 0.01 | 1.16 | 5.17 | 6.17 | 7.18 | 9.10 | 8.18 | 7.16 | 6.17 | 5.15 | 3.14 | 2.18 | 1.18 | 0.99 |
| 0.10 | 1.15 | 2.17 | 3.72 | -1.71 | 6.39 | 5.16 | 3.13 | 1.89 | 0.90 | 0.91 | 0.11 | -0.20 | 1.89 | 2.84 | 3.92 | 2.99 | 2.21 | 1.73 | ? |

2,99 2,21 1,73 ?
<table>
<thead>
<tr>
<th>1.34</th>
<th>1.78</th>
<th>2.56</th>
<th>1.88</th>
<th>0.57</th>
<th>-1.25</th>
<th>0.19</th>
<th>3.18</th>
<th>4.13</th>
<th>2.22</th>
<th>1.34</th>
<th>0.26</th>
<th>-1.90</th>
<th>-2.29</th>
<th>0.88</th>
<th>1.28</th>
<th>3.31</th>
<th>4.12</th>
<th>5.15</th>
<th>3.31</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.27</td>
<td>2.89</td>
<td>2.12</td>
<td>1.78</td>
<td>2.67</td>
<td>-3.16</td>
<td>0.01</td>
<td>1.16</td>
<td>5.17</td>
<td>6.17</td>
<td>7.18</td>
<td>9.10</td>
<td>8.18</td>
<td>7.16</td>
<td>6.17</td>
<td>5.15</td>
<td>3.14</td>
<td>2.18</td>
<td>1.18</td>
<td>0.99</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.10</td>
<td>1.15</td>
<td>2.17</td>
<td>3.72</td>
<td>-1.71</td>
<td>6.39</td>
<td>5.16</td>
<td>3.13</td>
<td>1.89</td>
<td>0.90</td>
<td>0.91</td>
<td>0.11</td>
<td>-0.20</td>
<td>1.89</td>
<td>2.84</td>
<td>3.92</td>
<td>2.99</td>
<td>2.21</td>
<td>1.73</td>
<td></td>
</tr>
</tbody>
</table>

Benoît Patra (Lokad - Université Paris VI)
<p>| | | | | | | | | | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1,34</td>
<td>1,78</td>
<td>2,56</td>
<td>1,88</td>
<td>0,57</td>
<td>−1,25</td>
<td>0,19</td>
<td>3,18</td>
<td>4,13</td>
<td>2,22</td>
<td>1,34</td>
<td>0,26</td>
<td>−1,90</td>
<td>−2,29</td>
<td>0,88</td>
<td>1,28</td>
</tr>
<tr>
<td>2,27</td>
<td>2,89</td>
<td>2,12</td>
<td>1,78</td>
<td>2,67</td>
<td>−3,16</td>
<td>0,01</td>
<td>1,16</td>
<td>5,17</td>
<td>6,17</td>
<td>7,18</td>
<td>9,10</td>
<td>8,18</td>
<td>7,16</td>
<td>6,17</td>
<td>5,15</td>
</tr>
<tr>
<td>0,10</td>
<td>1,15</td>
<td>2,17</td>
<td>3,72</td>
<td>−1,71</td>
<td>6,39</td>
<td>5,16</td>
<td>3,13</td>
<td>1,89</td>
<td>0,90</td>
<td>0,91</td>
<td>0,11</td>
<td>−0,20</td>
<td>1,89</td>
<td>2,84</td>
<td>3,92</td>
</tr>
</tbody>
</table>

2,99 2,21 1,73 ?
<table>
<thead>
<tr>
<th>1,34</th>
<th>1,78</th>
<th>2,56</th>
<th>1,88</th>
<th>0,57</th>
<th>−1,25</th>
<th>0,19</th>
<th>3,18</th>
<th>4,13</th>
<th>2,22</th>
<th>1,34</th>
<th>0,26</th>
<th>−1,90</th>
<th>−2,29</th>
<th>0,88</th>
<th>1,28</th>
<th>3,31</th>
<th>4,12</th>
<th>5,15</th>
<th>3,31</th>
</tr>
</thead>
<tbody>
<tr>
<td>2,27</td>
<td>2,89</td>
<td>2,12</td>
<td>1,78</td>
<td>2,67</td>
<td>−3,16</td>
<td>0,01</td>
<td>1,16</td>
<td>5,17</td>
<td>6,17</td>
<td>7,18</td>
<td>9,10</td>
<td>8,18</td>
<td>7,16</td>
<td>6,17</td>
<td>5,15</td>
<td>3,14</td>
<td>2,18</td>
<td>1,18</td>
<td>0,99</td>
</tr>
<tr>
<td>0,10</td>
<td>1,15</td>
<td>2,17</td>
<td>3,72</td>
<td>−1,71</td>
<td>6,39</td>
<td>5,16</td>
<td>3,13</td>
<td>1,89</td>
<td>0,90</td>
<td>0,91</td>
<td>0,11</td>
<td>−0,20</td>
<td>1,89</td>
<td>2,84</td>
<td>3,92</td>
<td>2,99</td>
<td>2,21</td>
<td>1,73</td>
<td></td>
</tr>
</tbody>
</table>

2,99 2,21 1,73

2,99 2,21 1,73

1,78 2,56 1,88 0,57

2,89 2,12 1,78 2,67

3,14 2,18 1,18 0,99

3,13 1,89 0,90 0,91

Benoît Patra (Lokad - Université Paris VI)
<table>
<thead>
<tr>
<th>1.34</th>
<th>1.78</th>
<th>2.56</th>
<th>1.88</th>
<th>0.57</th>
<th>-1.25</th>
<th>0.19</th>
<th>3.18</th>
<th>4.13</th>
<th>2.22</th>
<th>1.34</th>
<th>0.26</th>
<th>-1.90</th>
<th>-2.29</th>
<th>0.88</th>
<th>1.28</th>
<th>3.31</th>
<th>4.12</th>
<th>5.15</th>
<th>3.31</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.27</td>
<td>2.89</td>
<td>2.12</td>
<td>1.78</td>
<td>2.67</td>
<td>-3.16</td>
<td>0.01</td>
<td>1.16</td>
<td>5.17</td>
<td>6.17</td>
<td>7.18</td>
<td>9.10</td>
<td>8.18</td>
<td>7.16</td>
<td>6.17</td>
<td>5.15</td>
<td>3.14</td>
<td>2.18</td>
<td>1.18</td>
<td>0.99</td>
</tr>
<tr>
<td>0.10</td>
<td>1.15</td>
<td>2.17</td>
<td>3.72</td>
<td>-1.71</td>
<td>6.39</td>
<td>5.16</td>
<td>3.13</td>
<td>1.89</td>
<td>0.90</td>
<td>0.91</td>
<td>0.11</td>
<td>-0.20</td>
<td>1.89</td>
<td>2.84</td>
<td>3.92</td>
<td>2.99</td>
<td>2.21</td>
<td>1.73</td>
<td></td>
</tr>
</tbody>
</table>

- $? = \text{empirical quantile}$
Prediction and Aggregation.

Prediction of one expert.

\[ h_n^{(k,\ell)}(y_1^{n-1}) = \arg\min_{q \in \mathbb{R}} \sum_{\{t \in J_n^{(k,\ell)}\}} \rho_T(y_t - q). \]

[Can be easily computed by sorting the sample.]

Aggregated prediction of all experts.

\[ g_n(y_1^{n-1}) = \sum_{k,\ell=1}^{\infty} p_{k,\ell,n} h_n^{(k,\ell)}(y_1^{n-1}). \]

Where do the \( p_{k,\ell,n} \) come from?

Exponentially weight the experts based on their past performance.
Prediction and Aggregation.

**Prediction of one expert.**

\[ h_{n}^{(k,\ell)}(y_{1}^{n-1}) = \arg\min_{q \in \mathbb{R}} \sum_{\{t \in \mathcal{J}_{n}(k,\ell)\}} \rho_{\tau}(y_{t} - q). \]

[Can be easily computed by sorting the sample.]

**Aggregated prediction of all experts.**

\[ g_{n}(y_{1}^{n-1}) = \sum_{k,\ell=1}^{\infty} p_{k,\ell,n} h_{n}^{(k,\ell)}(y_{1}^{n-1}). \]

Where do the \( p_{k,\ell,n} \) come from?

Exponentially weight the experts based on their past performance.
Prediction and Aggregation.

Prediction of one expert.

\[ h_n^{(k,\ell)}(y_1^{n-1}) = \arg\min_{q \in \mathbb{R}} \sum_{\{t \in J_n^{(k,\ell)}\}} \rho_\tau(y_t - q). \]

[Can be easily computed by sorting the sample.]

Aggregated prediction of all experts.

\[ g_n(y_1^{n-1}) = \sum_{k,\ell=1}^{\infty} p_{k,\ell,n} h_n^{(k,\ell)}(y_1^{n-1}). \]

Where do the \( p_{k,\ell,n} \) come from?

Exponentially weight the experts based on their past performance.
Definitions.

- Let $\{q_{k,\ell}\}$ be a **probability distribution** over all pairs $(k, \ell)$ of positive integers such that $q_{k,\ell} > 0$ for all $(k, \ell)$.
- For $\eta_n > 0$, we define the **weights**

$$w_{k,\ell,n} = q_{k,\ell}e^{-\eta_n(n-1)L_{n-1}(h_n^{(k,\ell)})}.$$  

- We normalize these weights:

$$p_{k,\ell,n} = \frac{w_{k,\ell,n}}{\sum_{i,j=1}^{\infty} w_{i,j,n}}.$$  

Global prediction.

$$g_n(y_1^{n-1}) = \sum_{k,\ell=1}^{\infty} p_{k,\ell,n} h_n^{(k,\ell)}(y_1^{n-1}).$$
Definitions.

- Let \( \{q_{k,\ell}\} \) be a probability distribution over all pairs \((k, \ell)\) of positive integers such that \(q_{k,\ell} > 0\) for all \((k, \ell)\).

- For \( \eta_n > 0 \), we define the weights

\[
w_{k,\ell,n} = q_{k,\ell} e^{-\eta_n(n-1)L_{n-1}(h_n^{(k,\ell)})}.
\]

- We normalize these weights:

\[
p_{k,\ell,n} = \frac{w_{k,\ell,n}}{\sum_{i,j=1}^{\infty} w_{i,j,n}}.
\]

Global prediction.

\[
g_n(y_{1}^{n-1}) = \sum_{k,\ell=1}^{\infty} p_{k,\ell,n} h_n^{(k,\ell)}(y_{1}^{n-1}).
\]
Definitions.

- Let \( \{ q_{k,\ell} \} \) be a probability distribution over all pairs \( (k, \ell) \) of positive integers such that \( q_{k,\ell} > 0 \) for all \( (k, \ell) \).
- For \( \eta_n > 0 \), we define the weights
  \[
  w_{k,\ell,n} = q_{k,\ell} e^{-\eta_n(n-1)L_{n-1}(h_n^{(k,\ell)})}.
  \]
- We normalize these weights:
  \[
  p_{k,\ell,n} = \frac{w_{k,\ell,n}}{\sum_{i,j=1}^{\infty} w_{i,j,n}}.
  \]

Global prediction.

\[
 g_n(y_1^{n-1}) = \sum_{k,\ell=1}^{\infty} p_{k,\ell,n} h_n^{(k,\ell)}(y_1^{n-1}).
\]
Definitions.

- Let \( \{q_{k,\ell}\} \) be a probability distribution over all pairs \((k, \ell)\) of positive integers such that \( q_{k,\ell} > 0 \) for all \((k, \ell)\).
- For \( \eta_n > 0 \), we define the weights

\[
    w_{k,\ell,n} = q_{k,\ell} e^{-\eta_n(n-1)L_{n-1}(h_n^{(k,\ell)})}.
\]

- We normalize these weights:

\[
    p_{k,\ell,n} = \frac{w_{k,\ell,n}}{\sum_{i,j=1}^{\infty} w_{i,j,n}}.
\]

Global prediction.

\[
    g_n(y_1^{n-1}) = \sum_{k,\ell=1}^{\infty} p_{k,\ell,n} h_n^{(k,\ell)}(y_1^{n-1}).
\]
Theoretical Results.

**Theorem**

Let $\mathcal{C}$ be the class of all jointly **stationary** and **ergodic** processes $\{Y_n\}_{-\infty}^{\infty}$ such that $\mathbb{E}\{Y_0^2\} < \infty$ and $F_{Y_0|Y_{-\infty}^{-1}}$ is a.s. increasing.

Then the nearest neighbor quantile forecasting strategy is **universally consistent** with respect to the class $\mathcal{C}$, that is, for all process $Y \in \mathcal{C}$

$$\lim_{n \to \infty} L_n(g) = L^* \text{ almost surely.}$$
Theoretical Results.

**Theorem**

Let $C$ be the class of all jointly **stationary** and **ergodic** processes $\{Y_n\}_{-\infty}^{\infty}$ such that $\mathbb{E}\{Y_0^2\} < \infty$ and $F_{Y_0|Y_{-\infty}^{-1}}$ is a.s. increasing.

Then the **nearest neighbor quantile forecasting strategy** is **universally consistent** with respect to the class $C$, that is, for all process $Y \in C$

$$\lim_{n \to \infty} L_n(g) = L^* \quad \text{almost surely.}$$
Call center data set.

- Daily call volumes entering a call center.
- Long series between 382 and 826 time values. 21 series.
Future outcome predictions results.

\[ \tau = 0.5 \text{ median base forecaster: robustness.} \]

<table>
<thead>
<tr>
<th>Model Name</th>
<th>Avg Abs Error</th>
<th>Avg Sqr Error</th>
<th>Mape (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AR(7)</td>
<td>65.80</td>
<td>9738</td>
<td>31.6</td>
</tr>
<tr>
<td>QAR(8)_{0.5}</td>
<td>57.8</td>
<td>9594</td>
<td>24.9</td>
</tr>
<tr>
<td>DayOfTheWeekMean</td>
<td>53.95</td>
<td>7099</td>
<td>22.8</td>
</tr>
<tr>
<td>HoltWinters</td>
<td>49.84</td>
<td>6025</td>
<td>21.5</td>
</tr>
<tr>
<td>QuantileExpertMixture_{0.5}</td>
<td>48.1</td>
<td>5731</td>
<td>21.6</td>
</tr>
<tr>
<td>MeanExpertMixture</td>
<td>52.37</td>
<td>6536</td>
<td>22.3</td>
</tr>
<tr>
<td>MA</td>
<td>179</td>
<td>62448</td>
<td>0.52</td>
</tr>
</tbody>
</table>

**Figure:** Forecasting future outcomes.
Quantile forecasting.

<table>
<thead>
<tr>
<th>Model Name</th>
<th>PinBall Loss (0.1)</th>
<th>Ramp Loss</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\text{QuantileExpertMixture}_{0.1}$</td>
<td>13.71</td>
<td>0.80</td>
</tr>
<tr>
<td>$\text{QAR}(7)_{0.1}$</td>
<td>13.22</td>
<td>0.88</td>
</tr>
</tbody>
</table>

Figure: $\tau = 0.1$

<table>
<thead>
<tr>
<th>Model Name</th>
<th>PinBall Loss (0.9)</th>
<th>Ramp Loss</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\text{QuantileExpertMixture}_{0.9}$</td>
<td>12.27</td>
<td>0.07</td>
</tr>
<tr>
<td>$\text{QAR}(7)_{0.9}$</td>
<td>19.31</td>
<td>0.07</td>
</tr>
</tbody>
</table>

Figure: $\tau = 0.9$
Quantile forecasting.

Table:

<table>
<thead>
<tr>
<th>Model Name</th>
<th>PinBall Loss (0.1)</th>
<th>Ramp Loss</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\text{QuantileExpertMixture}_{0.1})</td>
<td>13.71</td>
<td>0.80</td>
</tr>
<tr>
<td>(\text{QAR(7)}_{0.1})</td>
<td>13.22</td>
<td>0.88</td>
</tr>
</tbody>
</table>

Figure: \(\tau = 0.1\)

Table:

<table>
<thead>
<tr>
<th>Model Name</th>
<th>PinBall Loss (0.9)</th>
<th>Ramp Loss</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\text{QuantileExpertMixture}_{0.9})</td>
<td>12.27</td>
<td>0.07</td>
</tr>
<tr>
<td>(\text{QAR(7)}_{0.9})</td>
<td>19.31</td>
<td>0.07</td>
</tr>
</tbody>
</table>

Figure: \(\tau = 0.9\)
Questions?