# Non-parametric estimation of conditional quantile functions for AR(1)-ARCH(1) processes

LEMA L. SEKNEWNA[1] PETER N. MWITA[2] BENJAMIN K. MUEMA[3]

**Abstract**

In this paper, non-parametric estimations of conditional quantile functions for time series with AR(1)-ARCH(1) scheme, represented by $X_t = \alpha(Z_t) + \varpi(Z_t)\varepsilon_t$ are carried out. An algorithm to estimating two quantile functions robustly is proposed and a use of a prediction method for non-parametric conditional quantile regression was adopted to deal with the problem of boundary effects due to outliers. Our estimations are proven to be more accurate than the existing and very simple to compute. An overview of the data generating process is given to ascerntain stationaruty of the process. All the estimations were based on the quantile regression method by Koenker and Zaho using the minimization of the conditional expectation of a loss function.

**Keywords & phrases:** Prediction; Conditional Quantile; Convergence; Kernel Distribution Estimation; Quantile Autoregression; Heteroscedasticity; Time Series; Inversion.

**2010 MS Classification**: 62G07, 62G05, 62G20, 62G35, 62G30

## 1 Introduction

In many regression methods, its usually about finding a linear or curvilinear relationship based on the scatter plot. Most regression methods estimate the average (mean) value of the response variable. Some $z$-$x$ scatter plots do not obey this dictatorship due to influential points also known as outliers. Financial and insurance data among others have significant variability and are in some cases known as heavy-tailed data [**?** ]. Those data possess isolated points (from the cloud) that distort any attempt to make a simple linear or other average-based regression. This is one of the reasons why many robust methods are being developed in both parametric and non-parametric ways. Robust because they aim to get rid of being influenced by extreme values. This is the case in methods as LAD (Least Absolute Deviations) which estimate the median or 1/2-quantile value of the response variable (see [**?** ]). Conditional quantile regression as developped in [**?** ] is more general and gives a more general description of the response variable at each level in $(0, 1)$. The local polynomial regression method, mostly used for non-parametric estimations, is robust but is still influenced by abnormally far-off points

[1]Department of Mathematics, Pan African University Institute for Basic Sciences, Technology and Innovation – Jomo Kenyatta University of Agriculture and Technology,P.O Box 62000-00200, Nairobi, Kenya.

[2]Department of Mathematics, Machakos University, P.O Box 136 - 90100, Machakos, KENYA.

[3]Department of Statistics and Actuarial Sciences, Jomo Kenyatta University of Agriculture and Technology, P.O Box 62000-00200, Nairobi, Kenya.

at boundaries. Outliers pull the curve toward them in places where there are few amounts of points. [? ] devised a method to perform the analysis without deleting them by filling the gap between the dense cloud and the very distant points by adding pseudo-points before making the non-parametric estimation of the probability density function. Our approach, in this paper, gives absolute robustness to these non-parametric methods estimates by solving the problem of outliers, smoothing the estimators and giving the possibility in forecasting. We base our estimations on the Nadaraya [? ] - Watson [? ] (NW) method which is a particular case of local polynomial regression. The method consists of detecting points likely to change the behavior of the curves towards the borders by using the method of Tukey then making an estimation of the quantile as discussed in [? ] then reintegrating the ouliers by predicting their response variable by $k$-NN algorithm. The latter is a data mining tool with predictive power from observations using distance or similarity. Prediction is possible when estimates are smooth. We performed a two step-estimation which consist of estimating the quantile location shift or the QAR (Quantile AutoRegressive). After smoothing it and predicting the response for the ouliers (omitted in the first place), the CSF (Conditional scale function) is estimated from the residuals. Specific assumptions, also found in literature, are made to ascertain the consistency of ours estimations. The data generating process is discussed in section 2. The combination of smoothing method and the ouliers handling reduce the bias of the estimate compared to the results in [? ]. To illustrate that, we simulated identical processes in terms of parameters, then obtained estimates from the processes and computed the quadratic errors. These errors are very small and confirm that our estimates are accurate. In section 4, we discuss the empirical estimation of the conditional distribution function and its inverse. Our results can be used in finance in calculating CVaR (Conditional Value-at-Risk), expected shortfall, etc. Considering a Quantile Autoregressive model,

$$X_t = \alpha_\tau(Z_t) + u_t, \quad t = 1, 2, \ldots \tag{1.1}$$

where $\alpha_\tau(Z_t)$ is the $\tau^{th}$ Conditional Quantile Function of $X_t$ given $Z_t$ and the innovation $u_t$ are assumed to be independent and identically distributed with zero $\tau^{th}$ quantile and constant scale function, see [? ]. Rough kernel estimators of $\alpha_\tau(z)$ and $\varpi_\tau(z)$ were derived and their consistencies proven in [? ]. To improve the accuracy of the estimations, a bootstrap kernel estimator of $\alpha_\tau(Z_t)$ was determined and shown to be consistent, see [? ]. This paper extends [? ] by assuming that the innovations follow Quantile Autoregressive Conditional Heteroscedastic process similar to Autoregressive-Quantile Autoregressive Conditional Heteroscedastic process proposed in [? ]:

$$X_t = \alpha_\tau(Z_t) + \varpi_\tau(Z_t)\varepsilon_t, \quad t = 1, 2, \ldots \tag{1.2}$$

where $\alpha_\tau(Z_t)$ is the conditional $\theta$-quantile function of $X_t$ given $Z_t$; $\varpi_\tau(Z_t)$ is a conditional scale function at $\tau$-level and $\varepsilon_t$ is independent and identically distributed (i.i.d.) error with zero $\tau$-quantile and unit scale. The function $\varpi_\tau(Z_t)$ can be expressed as

$$\varpi_\tau(Z_t) = \lambda\varpi(Z_t) \tag{1.3}$$

where $\varpi(Z_t)$ is the so called volatility found in [? ] and [? ] which are some key references on Engle's ARCH models and $\lambda$ is a positive constant depending on $\tau$ (see [? ]). An example of this kind of function is Autoregressive - Generalized Autoregressive Conditional Heteroscedastic AR(1)-GARCH(1,1),

$$X_t = \alpha_t + \varpi_t e_t, \; t = 1, 2, \ldots, \tag{1.4}$$

where

$$\alpha_t = \mu + \delta X_{t-1} \tag{1.5}$$
$$u_t = \varpi e_t \tag{1.6}$$
$$\varpi_t = \sqrt{w + \alpha u_{t-1}^2 + \beta \varpi_{t-1}^2} \tag{1.7}$$
$$e_t \sim \mathcal{N}(0,1), \text{ independent of } X_{t-1} \tag{1.8}$$

and $\mu \in (-\infty, \infty), |\delta| < 1,\ \beta > 0, \alpha > 0, w > 0,\ \alpha + \beta < 1$. Note that $\alpha_t$ may also be an ARMA (see [? ]). The specifications for model (1.4) are given in section 2.4.

Considering other financial time series models, the model (1.1) can be seen as a robust generalization of AR-ARCH- models, introduced in [? ], and their non-parametric generalizations reviewed by [? ]. For instance, consider a financial time series model of AR($p$)-ARCH($p$)-type,

$$X_t = \alpha(Z_t) + \varpi(Z_t)e_t,\ t = 1, 2, \dots \tag{1.9}$$

Where $Z_t = (X_{t-1}, X_{t-2}, \cdots, X_{t-p})$, $\alpha(\cdot)$ and $\varpi(\cdot)$ arbitrary functions representing, respectively, the conditional mean and conditional variance of the process.

A partitioned stationary $\alpha$-mixed time series $(X_t, Z_t)$, where the $X_t \in \mathbb{R}$ and the variate $Z_t \in \mathbb{R}^d$ are respectively $\mathcal{A}_t$-measurable and $\mathcal{A}_{t-1}$-measurable is considered. For some $\tau \in (0, 1)$, the conditional $\tau$-quantile of $X_t$ given the past $F_{t-1}$ assumed to be determined by $Z_t$ is estimated. For simplicity, we assume that $Z_t = X_{t-1} \in \mathbb{R}$ throughout the rest of the discussion.

# 2 Model definition

**Definition 1.** *A process is said to be weakly stationary, if its first and second moments are time invariant. Meaning that*

$$\mathrm{E}[X_t] = \mathrm{E}[X_{t-1}] = \lambda < \infty, \qquad \forall t \tag{2.1}$$
$$\mathrm{V}(X_t) = \rho_0 < \infty, \qquad \forall t \text{ and} \tag{2.2}$$
$$\mathrm{Cov}(X_t, X_{t-k}) = \rho_k, \qquad \forall t, \forall k. \tag{2.3}$$

*The third property only depends on the difference $t - (t - k)$.*

In the next section, we discuss the properties of the model AR(1)-ARCH(1) that will be simulated for the application of our findings.

## 2.1 AR(1) process

Recall that the process of application or to be simulated is a combination of two processes. The first is the AR(1) represented by

$$X_t = \mu + \delta X_{t-1} + e_t \tag{2.4}$$

where $\mu \in \mathbb{R}$ is a constant and $e_t$ is white noise with mean 0, constant variance $\sigma_e^2$ and is independent of the lagged value $X_{t-1}$. This model represents some outputs, in financial time series for instance, that depend on their own previous values and an innovation term (stochastic term)

**Theorem 1.** *The AR(1) process is stationary and ergodic for* $|\delta| < 1$.

*Proof.* Using the definition 1, we specify the parameter that yield the stationarity of the AR(1) process.

$$\operatorname{E}[X_t] = \mu + \delta \operatorname{E}[X_{t-1}] + 0 \tag{2.5}$$
$$\lambda = \mu + \delta\lambda \tag{2.6}$$
$$= \frac{\mu}{1-\delta} \tag{2.7}$$

and

$$\operatorname{V}(X_t) = 0 + \operatorname{V}(\delta X_{t-1} + e_t) \tag{2.8}$$
$$\rho_0 = \delta^2 \operatorname{V}(X_{t-1}) + \operatorname{V}(e_t) + 2\underbrace{\operatorname{Cov}(X_{t-1}, e_t)}_{=0} \tag{2.9}$$
$$\rho_0 = \delta^2 \rho_0 + \sigma_e^2 \tag{2.10}$$
$$\rho_0 = \frac{\sigma_e^2}{1-\delta^2} \tag{2.11}$$

We calculate the covariance, for $k = 1$, as

$$\operatorname{Cov}(X_t, X_{t-1}) = \operatorname{E}[X_t X_{t-1}] - \operatorname{E}[X_t]\operatorname{E}[X_{t-1}] \tag{2.12}$$
$$\rho_1 = \operatorname{E}[\mu X_{t-1} + \delta X_{t-1}^2 + e_t X_{t-1}] - \frac{\mu^2}{(1-\delta)^2} \tag{2.13}$$
$$= \frac{\mu^2}{1-\delta} + \delta \operatorname{E}[X_t^2] - \frac{\mu^2}{(1-\delta)^2} \tag{2.14}$$
$$= \frac{-\mu^2\delta}{(1-\delta)^2} + \delta\left(\operatorname{V}(X_t) + \left(\operatorname{E}[X_t]\right)^2\right) \tag{2.15}$$
$$= \frac{-\mu^2\delta}{(1-\delta)^2} + \delta\left(\frac{\sigma_e^2}{1-\delta^2} + \frac{\mu^2}{(1-\delta)^2}\right) \tag{2.16}$$
$$= \delta\frac{\sigma_e^2}{1-\delta^2} \tag{2.17}$$

Now, for $k = 2$ and using the properties of the Covariance, we have

$$\operatorname{Cov}(X_t, X_{t-2}) = \operatorname{Cov}(\mu + \delta X_{t-1} + e_t, X_{t-2}) \tag{2.18}$$
$$\rho_2 = \operatorname{Cov}(\mu, X_{t-2}) + \delta \operatorname{Cov}(X_{t-1}, X_{t-2}) + \operatorname{Cov}(e_t, X_{t-2}) \tag{2.19}$$
$$= 0 + \delta\rho_1 + 0 \tag{2.20}$$
$$= \delta^2\frac{\sigma_e^2}{1-\delta^2} \tag{2.21}$$

We conclude that

$$\operatorname{Cov}(X_t, X_{t-k}) = \rho_k = \delta^k \frac{\sigma_e^2}{1-\delta^2} \tag{2.22}$$

□

## 2.2 ARCH(1) process

As the AR(1) models the outputs from the previous ones, the ARCH(1) is the modelization of the actual innovation as function of the previous ones too. ARCH-based process are being utilized in most of the current time series analysis in finance, economics, etc because they model the volatility. An ARCH(1) is depicted by

$$\varepsilon_t = \varpi e_t, \tag{2.23}$$

$$\varpi = \left(\omega + \alpha\varepsilon_{t-1}^2\right)^{\frac{1}{2}}, \qquad t = 1, 2, \dots \tag{2.24}$$

with the conditions $\omega > 0$, $\alpha < 1$ and $e_t$ i.i.d with zero mean and variance 1 and independent to $\varepsilon_{t-1}$. There conditions allow the data generation process to be stationary. To show it, we calculate the following statistics:

$$\begin{aligned}
\mathrm{E}[\varepsilon_t] &= \mathrm{E}\left[\left(\omega + \alpha\varepsilon_{t-1}^2\right)^{\frac{1}{2}} e_t\right] && (2.25)\\
&= \mathrm{E}\left[\left(\omega + \alpha\varepsilon_{t-1}^2\right)^{\frac{1}{2}}\right] \times \underbrace{\mathrm{E}\left[e_t\right]}_{=0} && (2.26)\\
&= 0. && (2.27)
\end{aligned}$$

Let's also introduce the conditional statistics that will enable the calculation the variance of the process.

### 2.2.1 Conditional expectation

The conditional expectation of the ARCH(1) process is

$$\begin{aligned}
\mathrm{E}\left[\varepsilon_t \mid \varepsilon_{t-1}\right] &= \mathrm{E}\left[\left(\omega + \alpha\varepsilon_{t-1}^2\right)^{\frac{1}{2}} e_t \mid \varepsilon_{t-1}\right] && (2.28)\\
&= \left(\omega + \alpha\varepsilon_{t-1}^2\right)^{\frac{1}{2}} \mathrm{E}[e_t \mid \varepsilon_{t-1}] && (2.29)\\
&= \left(\omega + \alpha\varepsilon_{t-1}^2\right)^{\frac{1}{2}} \mathrm{E}[e_t] && (2.30)\\
&= 0. && (2.31)
\end{aligned}$$

### 2.2.2 Conditional variance

$$\begin{aligned}
\mathrm{V}\left[\varepsilon_t \mid \varepsilon_{t-1}\right] &= \mathrm{V}\left[\left(\omega + \alpha\varepsilon_{t-1}^2\right)^{\frac{1}{2}} e_t \mid \varepsilon_{t-1}\right] && (2.32)\\
&= \mathrm{E}\left[\left(\omega + \alpha\varepsilon_{t-1}^2\right) e_t^2 \mid \varepsilon_{t-1}\right] && (2.33)\\
&= \left(\omega + \alpha\varepsilon_{t-1}^2\right) \mathrm{E}[e_t^2] && (2.34)\\
&= \omega + \alpha\varepsilon_{t-1}^2. && (2.35)
\end{aligned}$$

The variance of the process is therefore given by the law of total variance

$$\begin{aligned}
\mathrm{V}(\varepsilon_t) &= \mathrm{E}\left[\mathrm{V}(\varepsilon_t \mid \varepsilon_{t-1})\right] + \mathrm{V}\left(\mathrm{E}[\varepsilon_t \mid \varepsilon_{t-1}]\right) && (2.36)\\
&= \mathrm{E}\left[\omega + \alpha\varepsilon_{t-1}^2\right] && (2.37)\\
&= \omega + \alpha\,\mathrm{E}[\varepsilon_{t-1}^2] && (2.38)\\
&= \omega + \alpha\left(\mathrm{V}(\varepsilon_t) + \left(\mathrm{E}[\varepsilon_t]\right)^2\right) && (2.39)\\
&= \omega + \alpha\,\mathrm{V}(\varepsilon_t) && (2.40)\\
\mathrm{V}(\varepsilon_t) &= \frac{\omega}{1-\alpha}. && (2.41)
\end{aligned}$$

For this process, the covariance

$$\mathrm{Cov}(\varepsilon_t, \varepsilon_{t-k}) = 0 \qquad \forall k > 0. \tag{2.42}$$

## 2.3 GARCH(1,1) process

This process depends on both the previous innovation and the previous conditional variance. It's defined as

$$\begin{aligned}
&\varepsilon_t = \varpi e_t, && (2.43)\\
&\varpi = \left(\omega + \alpha\varepsilon_{t-1}^2 + \beta\varpi_{t-1}^2\right)^{\frac{1}{2}}, && (2.44)\\
&e_t \sim \mathcal{N}(0,1), \text{ independent of } \varepsilon_{t-1} \text{ and } \varpi_{t-1}, \qquad t = 1, 2, \ldots && (2.45)
\end{aligned}$$

Using the definition 1, we can show the specifications of the GARCH(1,1). We calculate, as in the previous section, the statistics

$$\begin{aligned}
\mathrm{E}[\varepsilon_t] &= \mathrm{E}\left[\left(\omega + \alpha\varepsilon_{t-1}^2 + \beta\varpi_{t-1}^2\right)^{\frac{1}{2}} e_t\right] && (2.46)\\
&= \mathrm{E}\left[\left(\omega + \alpha\varepsilon_{t-1}^2 + \beta\varpi_{t-1}^2\right)^{\frac{1}{2}}\right]\mathrm{E}\left[e_t\right] && (2.47)\\
&= 0. && (2.48)
\end{aligned}$$

The conditional expectation of the GARCH(1,1) process is given by

$$\begin{aligned}
\mathrm{E}[\varepsilon_t \mid \varepsilon_{t-1}] &= \mathrm{E}\left[\left(\omega + \alpha\varepsilon_{t-1}^2 + \beta\varpi_{t-1}^2\right)^{\frac{1}{2}} e_t \mid \varepsilon_{t-1}\right] && (2.49)\\
&= \mathrm{E}\left[\left(\omega + \alpha\varepsilon_{t-1}^2 + \beta\varpi_{t-1}^2\right)^{\frac{1}{2}}\right]\mathrm{E}\left[e_t \mid \varepsilon_{t-1}\right] && (2.50)\\
&= 0, && (2.51)
\end{aligned}$$

and the conditional variance

$$\begin{aligned}
\mathrm{V}(\varepsilon_t \mid \varepsilon_{t-1}) &= \mathrm{E}\left[\varepsilon_t^2 \mid \varepsilon_{t-1}\right] && (2.52)\\
&= \mathrm{E}\left[\left(\omega + \alpha\varepsilon_{t-1}^2 + \beta\varpi_{t-1}^2\right) e_t^2 \mid \varepsilon_{t-1}\right] && (2.53)\\
&= \mathrm{E}\left[\left(\omega + \alpha\varepsilon_{t-1}^2 + \beta\varpi_{t-1}^2\right) \mid \varepsilon_{t-1}\right]\mathrm{E}\left[e_t^2 \mid \varepsilon_{t-1}\right] && (2.54)\\
&= \omega + \alpha\varepsilon_{t-1}^2 + \beta\varpi_{t-1}^2. && (2.55)
\end{aligned}$$

The law of total variance yields

$$\mathrm{V}(\varepsilon_t) = \mathrm{E}[\varpi^2] + \mathrm{V}(0) \tag{2.56}$$
$$= \mathrm{E}\left[\omega + \alpha\varepsilon_{t-1}^2 + \beta\varpi_{t-1}^2\right] \tag{2.57}$$
$$= \omega + \alpha\,\mathrm{E}\left[\varepsilon_{t-1}^2\right] + \beta\,\mathrm{E}\left[\varpi_{t-1}^2\right] \tag{2.58}$$
$$= \omega + \alpha\,\mathrm{V}(\varepsilon_t) + \beta\,\mathrm{V}(\varepsilon_t) \tag{2.59}$$
$$\mathrm{V}(\varepsilon_t) = \frac{\omega}{1-\alpha-\beta}. \tag{2.60}$$

This variance is positive and finite for $\omega > 0$ and $\alpha + \beta < 1$.

## 2.4 AR(1)-GARCH(1,1)

A financial time series can be of this form which is function of the previous return and the previous volatility or innovation. Our data generation process will be of the form:

$$X_t = \alpha_t + u_t \tag{2.61}$$
$$\alpha_t = \mu + \delta X_{t-1} \tag{2.62}$$
$$u_t = \varpi e_t \tag{2.63}$$
$$\varpi_t = \left(\omega + \alpha X_{t-1}^2 + \beta\varpi_{t-1}^2\right)^{\frac{1}{2}} \tag{2.64}$$
$$e_t \sim \mathcal{N}(0,1), \text{ independent of } X_{t-1}. \tag{2.65}$$

Here, we also calculate the statistics using the definition 1 in order to show the conditions over the coefficients that ascertain the stationarity of the process. The first moment is given by

$$\mathrm{E}[X_t] = \mathrm{E}\left[\mu + \delta X_{t-1} + \left(\omega + \alpha u_{t-1}^2 + \beta\varpi_{t-1}^2\right)^{\frac{1}{2}} e_t\right] \tag{2.66}$$
$$= \mu + \delta\,\mathrm{E}[X_{t-1}] + \mathrm{E}\left[\left(\omega + \alpha u_{t-1}^2 + \beta\varpi_{t-1}^2\right)^{\frac{1}{2}}\right]\mathrm{E}[e_t] \tag{2.67}$$
$$= \mu + \delta\,\mathrm{E}[X_t] \tag{2.68}$$
$$\mathrm{E}[X_t] = \frac{\mu}{1-\delta}. \tag{2.69}$$

### 2.4.1 Conditional expectation

$$\mathrm{E}[X_t \mid X_{t-1}] = \mu + \delta X_{t-1} + \mathrm{E}\left[\left(\omega + \alpha u_{t-1}^2 + \beta\varpi_{t-1}^2\right)^{\frac{1}{2}} e_t \mid X_{t-1}\right] \tag{2.70}$$
$$= \mu + \delta X_{t-1}. \tag{2.71}$$

### 2.4.2 Conditional variance

$$\mathrm{V}(X_t \mid X_{t-1}) = \mathrm{E}[X_t^2 \mid X_{t-1}] - (\mu + \delta X_{t-1})^2 \tag{2.72}$$
$$= \mathrm{E}\left[\left(\omega + (\alpha e_{t-1}^2 + \beta)\varpi_{t-1}^2\right) \mid X_{t-1}\right] \times \mathrm{E}[e_t^2 \mid X_{t-1}] \tag{2.73}$$
$$= \omega + (\alpha+\beta)\,\mathrm{E}\left[\varpi_{t-1}^2 \mid X_{t-1}\right] \tag{2.74}$$

### 2.4.3 Law of total variance

$$
\begin{aligned}
\mathrm{V}(X_t) &= \mathrm{E}\left[\mathrm{V}(X_t \mid X_{t-1})\right] + \mathrm{V}\left(\mathrm{E}[X_t \mid X_{t-1}]\right) && (2.75)\\
&= \mathrm{E}\left[\omega + (\alpha+\beta)\,\mathrm{E}\left[\varpi_{t-1}^2 \mid X_{t-1}\right]\right] + \mathrm{V}\left(\mu + \delta X_{t-1}\right) && (2.76)\\
&= \omega + (\alpha+\beta)\,\mathrm{E}\left[\varpi_{t-1}^2\right] + \delta^2\,\mathrm{V}(X_t) && (2.77)\\
\left(1-\delta^2\right) V[X_t] &= \omega + (\alpha+\beta)\,\mathrm{E}\left[\varpi_{t-1}^2\right] && (2.78)
\end{aligned}
$$

We have

$$\mathrm{E}\left[\varpi_t^2\right] = \omega + (\alpha+\beta)\,\mathrm{E}\left[\varpi_{t-1}^2\right] \tag{2.79}$$

and for stationary, we'll assume the moments to be time-independent. That is,

$$\mathrm{E}\left[\varpi_t^2\right] = \frac{\omega}{1-\alpha-\beta} \tag{2.80}$$

Finally,

$$\mathrm{V}[X] = \frac{\omega}{\left(1-\delta^2\right)(1-\alpha-\beta)} \tag{2.81}$$

which is positive and finite for $\omega > 0, |\delta| < 1$ and $\alpha + \beta < 1$.

# 3 Simulation of AR(1)-ARCH(1) processes

All our estimations will take into account a data generated from an AR(1)-ARCH(1), a process as in the section 2.4 wherem the GARCH term $\beta = 0$. In order to graphically show how the curves behave in view of the variation of the coefficients satisfying the conditions and which do not (See Figure 3.1, 3.2, 3.3 and 3.4). The Figure 3.3 and Figure 3.4 show non-stationary process because the parameters do not satisfy the conditions discussed in the previous section.

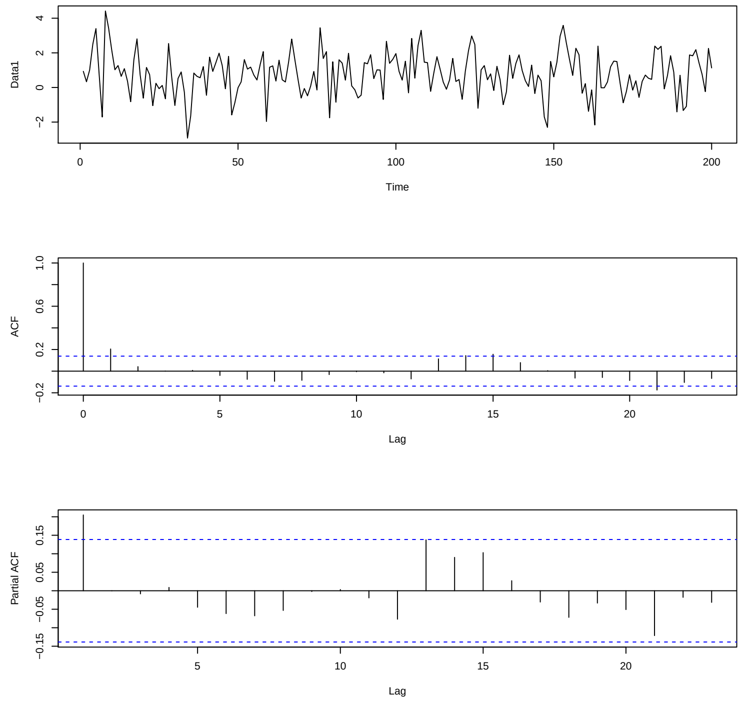


Figure 3.1: AR(1)-ARCH(1) process for $\mu = 0.5$, $\delta = 0.25$, $\omega = 1$, $\alpha = 0.35$

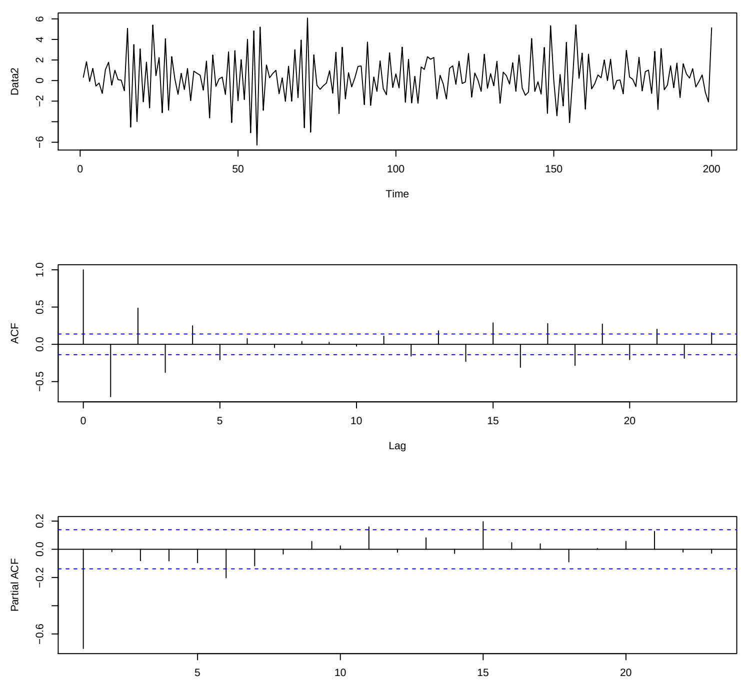


Figure 3.2: AR(1)-ARCH(1) process for $\mu = 0.5$, $\delta = -0.75$, $\omega = 1$, $\alpha = 0.5$

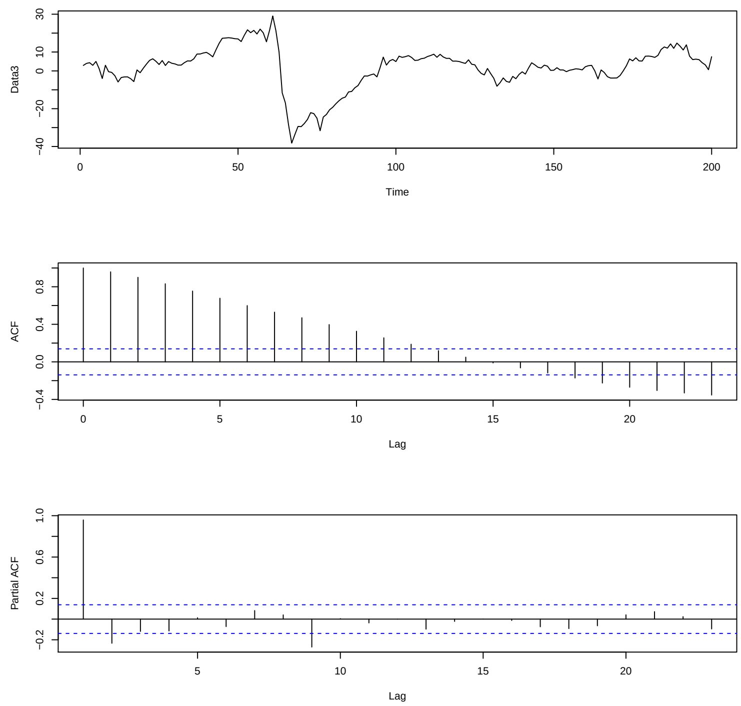


Figure 3.3: AR(1)-ARCH(1) process for $\mu = 0.5$, $\delta = 0.95$, $\omega = 1$, $\alpha = 1.2$

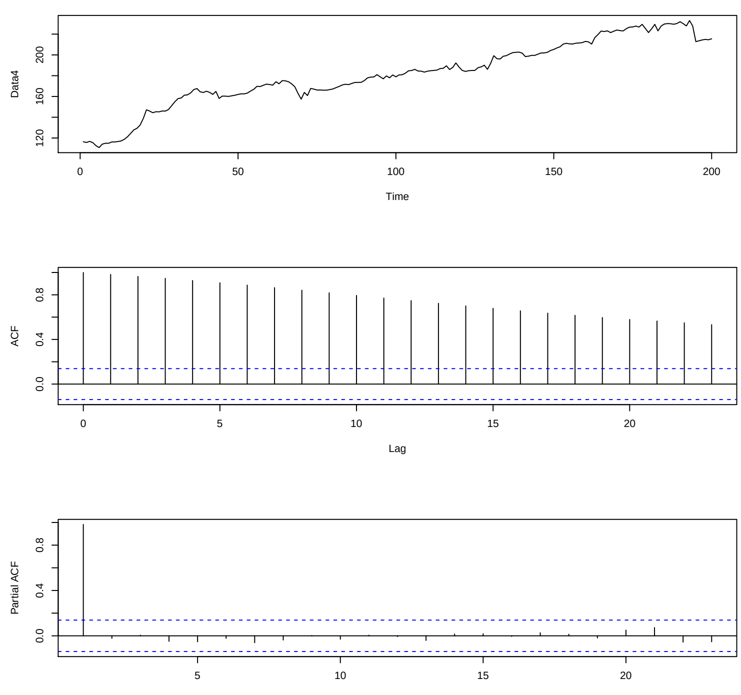


Figure 3.4: AR(1)-ARCH(1) process for $\mu = 0.5$, $\delta = 1$, $\omega = 1$, $\alpha = 1$

Now, having a clear information of the parameters that will come into play, we can simulate a stationary AR(1)-ARCH(1) (see Figure 3.1) process in order to apply our estimations that are discussed in the following section.

# 4 Estimation of quantile functions

To obtain the QAR-QARCH model from (1.1), we simply take its $\tau^{th}$ conditional quantile and we obtain:

$$Q_\tau(X_t \mid X_{t-1}) = \alpha_\tau(X_{t-1}) = \alpha(X_{t-1}) + \varpi(X_{t-1})q_\tau^e \tag{4.1}$$

where $q_\tau^e = F_e^{-1}(\tau)$ is the $\tau^{th}$ quantile of $\{e_t\}$. To make the reading less difficult, $X_{t-1}$ is changed to $Z_t$. Note that (4.1) is the estimation of the CVaR (Conditional Value-at-Risk) discussed in . Now, centering the response variable in (1.1) at its $\tau^{th}$-quantile in (4.1), we get:

$$X_t - \alpha_\tau(Z_t) = \varpi(Z_t)\,(e_t - q_\tau^e) \tag{4.2}$$

which is equivalent to the quantile autoregressive model:

$$X_t = \alpha_\tau(Z_t) + \varepsilon_\tau, \tag{4.3}$$

where $\varepsilon_\tau = \varpi(Z_t)\,(e_t - q_\tau^e)$ is 0 $\tau$-quantile, i.e, $Q_\tau(\varepsilon_\tau) = 0$.

We made the following assumptions:

**Assumption 1.** *The kernel function $K : \mathbb{R}^d \longrightarrow \mathbb{R}$ is symmetrical, non-negative and bounded satisfying $\int K(s)ds = 1$ with $\int_{\mathbb{R}^d} sK(s) = 0$.*

**Assumption 2.** *The process is strong mixing.*

The following definition, tells more about a strong mixing process.

**Definition 2 (strong mixing).** *A stationary process $X_t$ with $\sigma$-algebras $\mathcal{A}_t = \{X_j, -\infty < j \leq t\}$ and $\mathcal{A}^t = \{X_j, t \leq j < \infty\}$, $t = 1, \ldots, n$, is said to be strong mixing if*

$$\alpha(s) = \sup_{A \in \mathcal{A}_t,\, B \in \mathcal{A}^{t+s}} \left\{ \left| \mathrm{P}(A \cap B) - \mathrm{P}(A)\,\mathrm{P}(B) \right| \right\} \longrightarrow 0 \text{ as } n \longrightarrow \infty$$

**Assumption 3.** *The (positive) smoothing parameter is such that $b \to 0$, $nb \to \infty$ as $n \to \infty$.*

**Assumption 4.**
1. *$f(x, z)$ and $f(z)$ exist.*
2. *for fixed $(x, z)$, $0 < F(x|z) < 1$ and $f(z) > 0$ are continuous in the neighborhood of $z$ where the estimator is to be estimated.*
3. *The derivatives $F^{(j)}(x \mid z) = \frac{d^j F(x|z)}{dz^j}$ and $f^{(j)}(z) = \frac{d^j f(z)}{dz^j}$, for $j = 1, 2$, exist*
4. *$F(x \mid z)$ is a convex function in $x$ for fixed $z$.*
5. *The conditional density $f(x|z) = \frac{dF(x|z)}{dx}$ exists and is continuous in the neighborhood of $x$*
6. *$f\left(\alpha_\tau(z) \mid z\right) > 0$ and $f\left(\varpi_\tau(z) \mid z\right) > 0$*

## 4.1 Non-parametric QAR

**Theorem 2.** *Let* $\gamma_\tau(x,\mu) = \gamma_\tau(x-\mu) = \big(\tau - I(x-\mu \leq 0)\big)(x-\mu)$ *and* $(x,\sigma) \in \mathbb{R}^2$. *Then,* $\gamma_\tau$ *satisfies the Lipschitz continuity condition:*

$$\big|\gamma_\tau(x,\sigma) - \gamma_\tau(x,\sigma')\big| \leq M\big|\sigma - \sigma'\big|$$

*with the Lipschitz constant* $M = 1$ *and for all* $\sigma, \sigma'$.

***Proof of*** *Theorem 2.* Similar to the proof of Lemma 3.1 in [**?** , p .74-75] □

Consider the model (4.1) and the assumption made on the innovation $\varepsilon_\tau$. By definition, $\varepsilon_\tau$ is zero $\tau$-quantile meaning

$$\Pr(\varepsilon_\tau \leq 0) = \Pr(\varepsilon_\tau \leq 0 \mid Z_t) = \tau \tag{4.4}$$

and using (4.4), we have

$$\Pr(X_t \leq \alpha_\tau(Z) \mid Z_t) = \mathrm{E}\left[I\left(X_t \leq \alpha_\tau(Z_t)\right) \mid Z_t\right] = \tau \tag{4.5}$$

which is equivalent to $F\left(\alpha_\tau(Z_t) \mid Z_t\right) = \tau$. The conditional quantile function $\alpha_\tau$ minimizes the objective function $\mathrm{E}\left[\gamma_\tau\left(X_t, \alpha_\tau\right) \mid Z_t\right]$, i.e.

$$\alpha_\tau(z) = \arg\min_{\alpha_\tau} \mathrm{E}\left[\gamma_\tau\left(X, \alpha_\tau\right) \mid Z_t = z\right] \tag{4.6}$$

and is empirically given by

$$\hat{\alpha}_\tau(z) = \arg\min_{\alpha_\tau} \frac{1}{n}\sum_{t=1}^{n} K_b(Z_t - z)\gamma_\tau\left(X_t, \alpha_\tau\right) \tag{4.7}$$

Let's denote $\hat{\varphi}_{n,\tau} = \frac{1}{n}\sum_{t=1}^{n} K_b(Z_t - z)\gamma_\tau\left(X_t, \alpha_\tau\right)$. The zero of the equation $\frac{d}{d\alpha_\tau}\varphi_{n,\tau} = 0$ is

$$\hat{\alpha}_\tau(z) = \inf\big\{\mu : F(\mu \mid z) \geq \tau\big\} \equiv \hat{F}^{-1}(\tau \mid z) \tag{4.8}$$

where

$$\hat{F}(x \mid z) = \left[n\hat{f}(z)\right]^{-1}\sum_{t=1}^{n} K_b(Z_t - z) I\left(X_t \leq x\right) \tag{4.9}$$

Where $I(\cdot)$ is the indicator function which is 1 if the condition $X_t^* \leq x^*$ holds and 0 otherwise.

## 4.2 Empirical Conditional Distribution Function and its inverse

From the sequence $\{(X_t, Z_t)\}_{1 \leq t \leq n}$ of i.i.d. random variables, divide a span of our data into non-overlapping bins of the same size $z_1^* = \min(z_t) < z_2^* < \cdots < z_{n-1}^* < z_N^* = \max(z_t)$, $t = 1, 2, \ldots, n$ and compute the kernel matrix $K$ with elements given by

$$\left(k_{ij}\right)_{\substack{1 \leq i \leq N \\ 1 \leq j \leq n}} = K_b(z_i^* - Z_j) = \frac{1}{b}K\left(\frac{z_i^* - Z_j}{b}\right) \tag{4.10}$$

Where $K$ is the kernel density function (KDE) and $b$ is the smoothing parameter. The matrix of kernels is given by

$$M_K = \begin{pmatrix} K_b(z_1^* - Z_1) & K_b(z_1^* - Z_2) & \cdots & K_b(z_1^* - Z_n) \\ K_b(z_2^* - Z_1) & K_b(z_2^* - Z_2) & \cdots & K_b(z_2^* - Z_n) \\ \vdots & \vdots & \vdots & \vdots \\ K_b(z_N^* - Z_1) & K_b(z_N^* - Z_2) & \cdots & K_b(z_N^* - Z_n) \end{pmatrix} \tag{4.11}$$

The estimation of the empirical probability density function of $Z_t$ is given by

$$\hat{g}(z_i^*) = \frac{1}{n}\sum_{j=1}^{n} k_{ij} \tag{4.12}$$

and the matrix expression of $\hat{g}$

$$M_{\hat{g}} = \frac{1}{n} M_K \mathbb{1}_n, \quad \mathbb{1}_n = (1, 1, \ldots, 1)^T \in \mathbb{R}^n \tag{4.13}$$

$$= \frac{1}{n}\begin{pmatrix} K_b(z_1^* - Z_1) & K_b(z_1^* - Z_2) & \cdots & K_b(z_1^* - Z_n) \\ K_b(z_2^* - Z_1) & K_b(z_2^* - Z_2) & \cdots & K_b(z_2^* - Z_n) \\ \vdots & \vdots & \vdots & \vdots \\ K_b(z_N^* - Z_1) & K_b(z_N^* - Z_2) & \cdots & K_b(z_N^* - Z_n) \end{pmatrix}\begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} \tag{4.14}$$

$$= \frac{1}{n}\begin{pmatrix} K_b(z_1^* - Z_1) + K_b(z_1^* - Z_2) + \cdots + K_b(z_1^* - Z_n) \\ K_b(z_2^* - Z_1) + K_b(z_2^* - Z_2) + \cdots + K_b(z_2^* - Z_n) \\ \vdots \\ K_b(z_N^* - Z_1) + K_b(z_N^* - Z_2) + \cdots + K_b(z_N^* - Z_n) \end{pmatrix} \tag{4.15}$$

which is a vector of $N$ elements. We also introduce the indicator matrix $M_I$ with columns representing each $I(X_t \leq x)$ for fixed $x$ (for each column) and $t = 1, 2, \ldots, n$. The product of the kernel matrix $M_K$ and the matrix $M_I$ contains all the summations (also seen as joint probability density function at $X_t = x$ and $Z_t = z^*$).

$$\hat{f}(x, z^*) = \sum_{t=1}^{n} K_b(z^* - Z_t) I(X_t \leq x) \tag{4.16}$$

with matrix form $M_I$ for all fixed couple $(z^*, x) \in \mathbb{R}^2$.

$$M_I = \begin{pmatrix} I(x_1 \leq x_1) & I(x_1 \leq x_2) & \ldots & I(x_1 \leq x_n) \\ I(x_2 \leq x_1) & I(x_2 \leq x_2) & \ldots & I(x_2 \leq x_n) \\ \vdots & & & \\ I(x_n \leq x_1) & I(x_n \leq x_2) & \ldots & I(x_n \leq x_n) \end{pmatrix} \tag{4.17}$$

$$= \begin{pmatrix} 1 & I(x_1 \leq x_2) & \ldots & I(x_1 \leq x_n) \\ I(x_2 \leq x_1) & 1 & \ldots & I(x_2 \leq x_n) \\ \vdots & & & \\ I(x_n \leq x_1) & I(x_n \leq x_2) & \ldots & 1 \end{pmatrix} \tag{4.18}$$

The elements of $M_I$ are 1 where the inequalities are true and 0 otherwise. The matrix of the joint probability density function in (4.16) is

$$
\begin{aligned}
M_{\hat{f}} &= M_k M_I \\
&= \begin{pmatrix} K_b(z_1^* - Z_1) & K_b(z_1^* - Z_2) & \cdots & K_b(z_1^* - Z_n) \\ K_b(z_2^* - Z_1) & K_b(z_2^* - Z_2) & \cdots & K_b(z_2^* - Z_n) \\ \vdots & \vdots & \vdots & \vdots \\ K_b(z_N^* - Z_1) & K_b(z_N^* - Z_2) & \cdots & K_b(z_N^* - Z_n) \end{pmatrix} \times \\
&\quad \begin{pmatrix} 1 & I(x_1 \le x_2) & \dots & I(x_1 \le x_n) \\ I(x_2 \le x_1) & 1 & \dots & I(x_2 \le x_n) \\ \vdots & & & \\ I(x_n \le x_1) & I(x_n \le x_2) & \dots & 1 \end{pmatrix}
\end{aligned} \tag{4.19}
$$

and the one for conditional cumulative distribution functions (CCDF) is given by

$$
\left(F_{ji}\right)_{\substack{1 \le j \le n \\ 1 \le i \le N}} = \frac{\sum_{t=1}^{n} k_{it} \cdot I(X_t \le x_j)}{n \hat{g}(z_i^*)} \tag{4.20}
$$

with matrix form

$$
M_{\hat{F}} = M_K M_I / \left(M_K \mathbb{1}_{n \times n}\right) \tag{4.21}
$$

$$
= \begin{pmatrix} \dfrac{\sum\limits_{t=1}^{n} K_b(z_1^* - Z_t) I(X_t \le x_1)}{\sum\limits_{t=1}^{n} K_b(z_1^* - Z_t)} & \cdots & \dfrac{\sum\limits_{t=1}^{n} K_b(z_1^* - Z_t) I(X_t \le x_n)}{\sum\limits_{t=1}^{n} K_b(z_1^* - Z_t)} \\ \vdots & \vdots & \vdots \\ \dfrac{\sum\limits_{t=1}^{n} K_b(z_N^* - Z_t) I(X_t \le x_1)}{\sum\limits_{t=1}^{n} K_b(z_N^* - Z_t)} & \cdots & \dfrac{\sum\limits_{t=1}^{n} K_b(z_N^* - Z_t) I(X_t \le x_n)}{\sum\limits_{t=1}^{n} K_b(z_N^* - Z_t)} \end{pmatrix} \tag{4.22}
$$

Each element of the $(n \times N)$-matrix $M_{\hat{F}}$ is the computation of $\hat{F}(x_j \mid z_i^*)$. For each row $i$ of $F$, $1 \le i \le N$, we choose the minimum of $x_j$'s that satisfy $\hat{F}(x_j \mid z_i) \ge \tau$, $\tau \in (0, 1)$. This estimates the QAR or $\hat{F}^{-1}(\tau \mid z^*)$. We notice that the number of selected $x_j$'s will exactly be the number of bins.

## 4.3 $k$ Nearest Neighbor ($k$-NN) prediction

The prediction $\tilde{\alpha}_\tau(z)$ of a future value or any value $Z_{n+1} = z$ is easy in parametric regression once we have the estimated coefficients of a model. But in non-parametric regression, this prediction is somehow impossible. Recent research on this problem suggests methods more or less feasible for our type of estimation. There is the $k$NN ($k$ Nearest Neighbor)[**?** ] method which consists of finding the $k$ values, $\underline{z}_1^*, \dots, \underline{z}_k^*$ close to $z$. The requirement of this method is that the estimator $\alpha_\tau$ is to be smooth [**?** ][**?** ]. Unfortunately, the estimation of the QAR in (4.7) is not smooth and suffers from boundary issues. Having estimated $\hat{\alpha}_\tau(z_i^*)$ and the bin points $z_i^*$, $i = 1, \ldots, N$, thus, $\tilde{\alpha}_\tau(z)$ will be the average of $\hat{\alpha}_\tau(\underline{z}_1^*), \ldots, \hat{\alpha}_\tau(\underline{z}_k^*)$. In other words,

$$\tilde{\alpha}_\tau(z) = \frac{1}{k} \sum_{i=1}^{k} \hat{\alpha}_\tau(\underline{z}_i^*) \tag{4.23}$$

This approach is used to predict the values $\tilde{\alpha}_\tau(Z_t)$ which is a sequence of $n$ points. The figure 8.1 represents the prediction for the entire data (for instance, the daily returns) at $\tau = 0.25, 0.50, 0.75, 0.9$. In order to see if the prediction is accurate, the following error is calculated (the mean squared error of the difference between $\hat{\alpha}_\tau(z_i^*)$ and $\tilde{\alpha}_\tau(z_i^*)$ for bins $z_1^*, \ldots, z_N^*$)

$$\tilde{e}_p = \frac{1}{N} \sum_{i=1}^{N} \left(\hat{\alpha}_\tau(z_i^*) - \tilde{\alpha}_\tau(z_i^*)\right)^2 \tag{4.24}$$

The same prediction applies when we have the non-parametric estimation of the conditional scale function $\hat{\varpi}_\tau$.

## 4.4 Non-parametric QARCH

Considering that the QAR is already estimated, we have

$$Q_\tau \Big[\gamma_\tau \big(X_t - \alpha_\tau(Z_t)\big)\Big] = \varpi(Z_t) Q_\tau \left[\gamma_\tau \left(e_t - q_\tau^e\right)\right] \tag{4.25}$$

The ratio of $X_t - \alpha_\tau(Z_t)$ in (4.2) and the left part in (4.25) gives

$$\frac{X_t - \alpha_\tau(Z_t)}{Q_\tau \Big[\gamma_\tau \big(X_t - \alpha_\tau(Z_t)\big)\Big]} = \frac{e_t - q_\tau^e}{Q_\tau \left[\gamma_\tau \left(e_t - q_\tau^e\right)\right]} \tag{4.26}$$

This transformation leads to the QAR-QARCH model

$$X_t = \alpha_\tau(Z_t) + \varpi_\tau(Z_t)\eta_\tau \tag{4.27}$$

where $\varpi_\tau(Z_t) = Q_\tau \Big[\gamma_\tau \big(X_t - \alpha_\tau(Z_t)\big)\Big]$ and $\eta_\tau = \frac{e_t - q_\tau^e}{Q_\tau \left[\gamma_\tau \left(e_t - q_\tau^e\right)\right]}$ is zero $\tau$-quantile with unit scale. This property leads to the expression

$$\Pr\big(\gamma_\tau(\eta_\tau) \leq 1\big) = \Pr\big(\gamma_\tau(\eta_\tau) \leq 1 \mid Z\big) = \tau \tag{4.28}$$

This is identifiable to (4.5), if $X_t$ and $\alpha_\tau(Z_t)$ are replaced by $\gamma_\tau\big(X_t - \alpha_\tau(Z_t)\big)$ and $\varpi_\tau(Z_t)$ respectively. Thus, $\varpi_\tau(Z_t)$ minimizes $\mathrm{E}\Big[\gamma_\tau\Big(\gamma_\tau\big(X_t, \alpha_\tau(Z_t)\big), \varpi_\tau(Z_t)\Big) \mid Z_t\Big]$, i.e.

$$\varpi_\tau(Z_t) = \arg\min_{\varpi_\tau} \mathrm{E}\Big[\gamma_\tau\big(X_t^*, \varpi_\tau\big) \mid Z_t\Big] \tag{4.29}$$

or is empirically given by

$$\hat{\varpi}_\tau(Z_t) = \arg\min_{\varpi_\tau} \frac{1}{n} \sum_{t=1}^{n} K_b(Z_t - z)\gamma_\tau(X_t^*, \varpi_\tau) \tag{4.30}$$

where $X_t^* = \gamma_\tau\left(X_t, \alpha_\tau(Z_t)\right)$. Again, if we denote $\hat{\varphi}_{n,\tau} = \frac{1}{n}\sum_{t=1}^n K_b(Z_t - z)\gamma_\tau\left(X_t^*, \varpi_\tau\right)$, then $\frac{d\hat{\varphi}_{n,\tau}}{d\varpi_\tau} = 0$ has as solution

$$\hat{\varpi}_\tau(z) = \inf\left\{x^* \in \mathbb{R}_*^+ : \hat{F}\left(x^* \mid z\right) \geq \tau\right\} \equiv \hat{F}^{-1}(\tau \mid z) \tag{4.31}$$

with

$$\hat{F}\left(x^* \mid z\right) = \left[n\hat{f}(z)\right]^{-1}\sum_{t=1}^n K_b(Z_t - z)I(X_t^* \leq x^*) \tag{4.32}$$

We prove the consistency of our estimations with the following theorem

**Theorem 3.** *Suppose that the assumptions 1, 2, 3 and 4 hold. Then, $\hat{\alpha}_\tau$ and $\hat{\varpi}_\tau$ are asymptotically normal in distribution.*

***Proof of*** *theorem 3.* The proof is found in our previous work [**?** ]. □

# 5 Bias reduction

## 5.1 Outliers detection

Before estimating the conditional quantile function $\hat{\alpha}_\tau$, we first did the detection of the far-off points which are points outside the interval

$$\left[Q_1 - 3\times(Q_3 - Q_1), \quad Q_3 + 3\times(Q_3 - Q_1)\right]$$

where $Q_1$ and $Q_3$ are the first and the third quantiles of the sequence of random variables $Z_1, Z_2, \ldots, Z_n$.

## 5.2 Kernel smoother

The idea here is to regress the rough QAR estimation (without outliers) on the bins $z_1^*, z_2^*, \ldots, z_N^*$ using Nadaraya - Watson kernel regression. The resulting smoothed curve is used to predicted the today's QAR for $Z_t$'s, $t = 1, 2, \ldots, n$. Figure 5.2 shows the limits of the interval. It shows also the rough estimation of the QAR without removing the outliers (red curve) and the predicted smooth curve

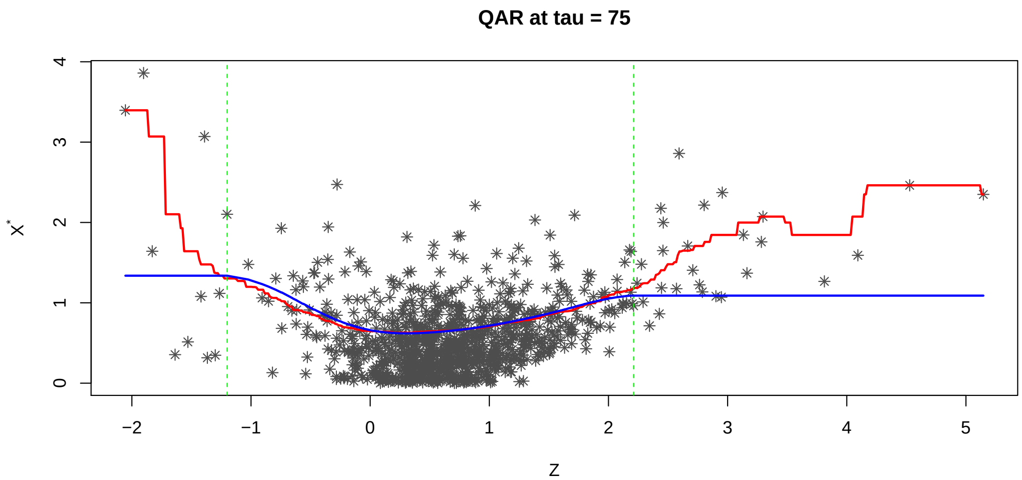


Figure 5.1: Rough (red) and smooth predicted (blue) QAR

# 6 Accuracy of estimations

In order to show the accuracy of our smooth estimators, we simulated (random) AR(1)-ARCH(1) process of size $m = 250, 500, 1000$ with same coefficients $\mu = 0.5$, $\delta = 0.3$, $\omega = 1$, $\alpha = 0.35$ and $e_t \sim \mathcal{N}(0,1)$. The following tables confirm the accuracy of the smooth estimations.

Table 6.1: MASE for $\tau = 0.25$

| $n$ | rough $\hat{\alpha}_{0.25}$ | smooth $\hat{\alpha}_{0.25}$ | rough $\hat{\varpi}_{0.25}$ | smooth $\hat{\varpi}_{0.25}$ |
|---|---|---|---|---|
| 250 | 1.13482 | 0.03078 | 0.03457 | 0.00075 |
| 500 | 0.94149 | 0.04128 | 0.04916 | 0.00075 |
| 1000 | 1.22881 | 0.00671 | 0.15645 | 0.00115 |

Table 6.2: MASE for $\tau = 0.50$ (median)

| $n$ | rough $\hat{\alpha}_{0.50}$ | smooth $\hat{\alpha}_{0.50}$ | rough $\hat{\varpi}_{0.50}$ | smooth $\hat{\varpi}_{0.50}$ |
|---|---|---|---|---|
| 250 | 0.6184 | 0.01963 | 0.08938 | 0.00401 |
| 500 | 1.21301 | 0.00873 | 0.3448 | 0.00526 |
| 1000 | 1.54507 | 0.0091 | 0.3595 | 0.00816 |

Table 6.3: MASE for $\tau = 0.75$

| $n$ | rough $\hat{\alpha}_{0.75}$ | smooth $\hat{\alpha}_{0.75}$ | rough $\hat{\varpi}_{0.75}$ | smooth $\hat{\varpi}_{0.75}$ |
|---|---|---|---|---|
| 250 | 1.88628 | 0.03351 | 1.36976 | 0.0126 |
| 500 | 0.39451 | 0.02664 | 0.69214 | 0.02616 |
| 1000 | 1.28018 | 0.01356 | 1.21384 | 0.02367 |

Table 6.4: MASE for $\tau = 0.90$

| $n$ | rough $\hat{\alpha}_{0.90}$ | smooth $\hat{\alpha}_{0.90}$ | rough $\hat{\varpi}_{0.90}$ | smooth $\hat{\varpi}_{0.90}$ |
|---|---|---|---|---|
| 250 | 0.66136 | 0.222 | 0.62655 | 0.17793 |
| 500 | 0.99836 | 0.12574 | 1.09674 | 0.27295 |
| 1000 | 1.76097 | 0.07349 | 1.47431 | 0.1794 |

# 7 Quantile error

From our previous paper [**?** ], we showed the asymptotic properties of the conditional scale function estimate through inversion of the conditional CCDF as in (4.32) with the assumption that the quantile location shift $\alpha_\tau$ is zero. The properties for the QAR estimate are the same given that the two CCDFs in (4.9) and (4.32) differ respectively in the conditional part $I(X_t \leq x)$ and $I(X_t^* \leq x^*)$ only. Thus, assuming we have estimated the two components using the prediction method, the quantile error $\eta_\tau$ can be estimate as

$$\hat{\eta}_\tau = \frac{X_t - \tilde{\alpha}_\tau(Z_t)}{\tilde{\varpi}_\tau(Z_t)} \tag{7.1}$$

and should verify the conditions (4.4) and (4.28). Moreover, if the conditions hold, then the estimators are accurate. From our simulation, the estimations seem to be accurate for quantile $\tau = 0.75$ (see Table 7.1).

Table 7.1: Summary of quantile errors

| $\tau$ | Min. | 1st Qu. | Med | Mean | 3rd Qu. | Max. | $\Pr(\eta_\tau \leq 0)$ | $\Pr(\eta^*_\tau \leq 1)$ |
|---|---|---|---|---|---|---|---|---|
| 0.25 | -47.18 | 0.03 | 3.00 | 3.53 | 6.48 | 61.04 | 0.25 | 0.42 |
| 0.50 | -16.78 | -1.46 | 0.02 | 0.17 | 1.61 | 22.99 | 0.50 | 0.62 |
| 0.75 | -16.61 | -2.86 | -1.49 | -1.32 | 0.06 | 21.53 | 0.74 | 0.74 |
| 0.90 | -16.62 | -2.79 | -1.95 | -1.92 | -1.09 | 12.96 | 0.89 | 0.96 |

where $\eta^*_\tau = \gamma_\tau(\eta_\tau)$.

# 8 Simulation study

The figure 8.1 represents the superposition of the process and the estimated $\tilde{\alpha}_\tau(z)$ using the $k$NN prediction method. In fact, the non-parametric estimation of $\hat{\alpha}_\tau(z)$ was first carried out using the smoothed estimator along with the outliers detection using box-plot fences in order to correct the boundary issue (see [**?** ]). The comparison between $\hat{\alpha}_\tau(z)$ and the predicted $\tilde{\alpha}_\tau(z)$ for bins $z$ is represented by Figure 8.2. Note that the prediction error in (4.24) was evaluated to $10^{-6}$ and the Figure 8.2 illustrates it as well. The outliers detection technique and prediction give less weight to extreme points that are not considered in the first estimation, then are re-involved in the prediction. This made our estimations less sensitive to the boundaries (see Figure 8.2).

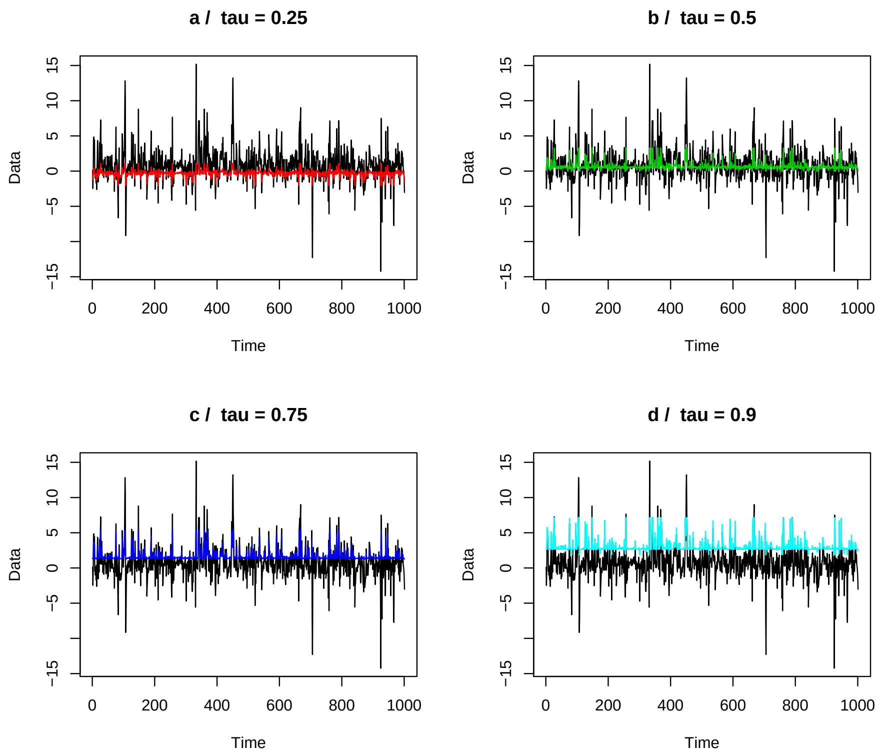


Figure 8.1: Predicted conditional quantile returns

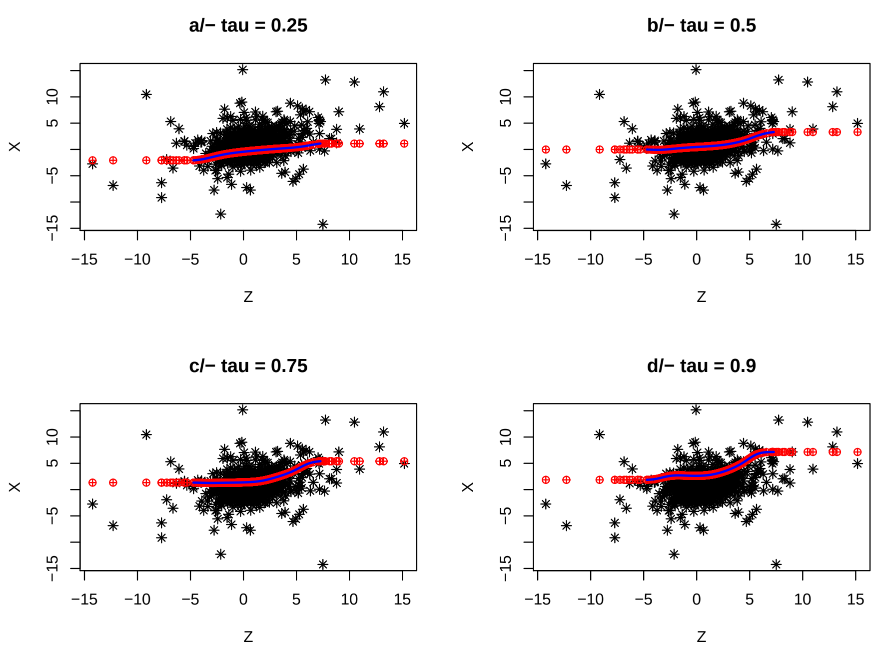


Figure 8.2: Graphical superposition of $\tilde{\alpha}_\tau(z)$ [red points] and $\hat{\alpha}_\tau(z)$ [blue curve]

# 9 Conclusion

In this paper, the problem of estimating the conditional scale function when the autoregressive part is not zero is carried out using Nadaraya-Watson kernel estimation and Quantile Autoregression method. The rough estimation of the QAR feels the boundaries and that increased the bias of the estimates. We were able to correct the boundary issue and showed the accuracy of our estimations. With the use of the $k$-NN method, it was possible to calculate the quantile error for each proportion $\tau \in (0, 1)$. The next step will the application of our approach on real data.

## Acknowledgment

The authors thank the PAU Institute for basic Sciences, Technology and Innovation for funding this research and all those who contributed to the realization of this article.