High Quantile Regression for Tail Dependent Time Series[1]

Ting Zhang

tingzhang@uga.edu

Department of Statistics
University of Georgia

June 27, 2023

BU-Keio-Tsinghua Workshop 2023, Boston, MA

[1]Research supported by NSF CAREER Award DMS-2131821.

What is quantile regression?

- Suppose we observe $(y_i, \boldsymbol{x}_i)$, $i = 1, \ldots, n$, where
  - y_i: response variable;
  - $\boldsymbol{x}_i$: a set of explanatory variables.
- Traditional linear regression:
  - concerns the effect of $\boldsymbol{x}_i$ on the mean of y_i;
  - regression coefficient can be estimated by least squares:

$$\check{\boldsymbol{\beta}}_n = \operatorname*{argmin}_{\boldsymbol{\eta} \in \mathbb{R}^p} \sum_{i=1}^{n} (y_i - \boldsymbol{x}_i^\top \boldsymbol{\eta})^2.$$

- Quantile regression (Koenker and Bassett, 1978, *Econometrica*):
  - concerns the effect of $\boldsymbol{x}_i$ on quantiles of y_i;
  - regression coefficient for the $(1-\alpha)$-th quantile can be estimated by

$$\tilde{\boldsymbol{\beta}}_n = \operatorname*{argmin}_{\boldsymbol{\eta} \in \mathbb{R}^p} \sum_{i=1}^{n} \phi_{1-\alpha}(y_i - \boldsymbol{x}_i^\top \boldsymbol{\eta}),$$

where $\phi_{1-\alpha}(y) = (1-\alpha)y^+ + \alpha(-y)^+$ is the check function.

Why high quantiles?

- Traditional quantile regression:
  - the quantile level $1 - \alpha$ is assumed to be fixed.
  - Examples:
    - median: $1 - \alpha = 0.5$;
    - first quartile: $1 - \alpha = 0.25$;
    - third quartile: $1 - \alpha = 0.75$;
    - ...
- Recent applications require the study of tail phenomena:
  - Abrevaya (2001, *Empirical Economics*): the analysis of low percentiles of the birthweight distribution;
  - Tsai (2012, *J. Int. Financial Mark. Inst. Money*): relationship between stock price index and exchange rates when exchange rates are extremely high or low;
  - Elsner et al. (2008, *Nature*): understanding temporal trends of strong tropical cyclones;
  - Rhines et al. (2017, *Journal of Climate*): trend analysis on temperatures of the coldest days in North America;
  - ...

How to accommodate high quantiles?

- Traditional quantile regression:

$$\tilde{\boldsymbol{\beta}}_n = \operatorname*{argmin}_{\boldsymbol{\eta}\in\mathbb{R}^p} \sum_{i=1}^{n} \phi_{1-\alpha}(y_i - \boldsymbol{x}_i^\top \boldsymbol{\eta}),$$

where the quantile level $1-\alpha$ is treated as fixed.

- High quantile regression:

$$\hat{\boldsymbol{\beta}}_n = \operatorname*{argmin}_{\boldsymbol{\eta}\in\mathbb{R}^p} \sum_{i=1}^{n} \phi_{1-\alpha_n}(y_i - \boldsymbol{x}_i^\top \boldsymbol{\eta}),$$

where the quantile level $1-\alpha_n \to 1$ as $n \to \infty$.
  - allows the quantile level to approach the unit as the sample size increases to capture the tail phenomena;
  - theoretical properties can be more challenging due to the double asymptotics;
  - convergence rate may no longer be the universal $n^{1/2}$ as in the conventional setting.

What has been done on high quantile regression?

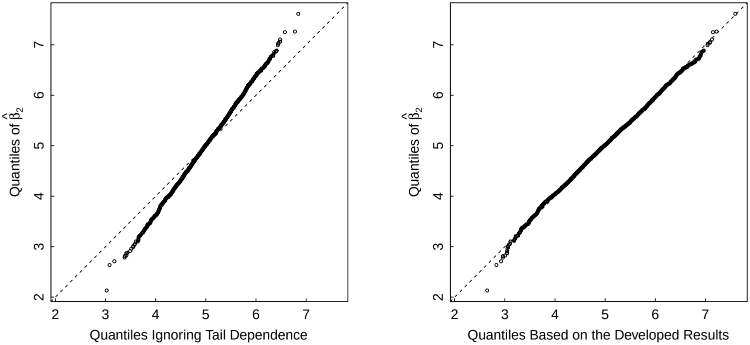
- Wang, Li and He (2012, *JASA*) commented that

  "To our knowledge, relatively little has been done for the estimation of high conditional quantiles."

- Existing results were mainly developed for independent data:
  - de Haan and Ferreira (2006);
  - Belloni and Chernozhukov (2011, *AoS*);
  - Wang, Li and He (2012, *JASA*);
  - Wang and Li (2013, *JASA*);
  - He, Cheng and Tong (2016);
  - Wang and Wang (2016, *Statistica Sinica*);
  - D'Haultfœuille, Maurel and Zhang (2018, *JoE*);
  - Zhang (2018, *AoS*);
  - ...

- What if the data exhibits dependence?
  - For time series data, dependence is the rule rather than the exception.

What if we simply ignore the dependence?



"Ignoring the dependence can lead to misleading p-values!"

	0.1	0.05	0.01
Ignoring the dependence	0.224	0.150	0.059
Developed results	0.087	0.043	0.007

What about high quantile regression under dependence?

- Chernozhukov (2005, *AoS*):
  - An influential work that considered the possibility of allowing time series data by using the strong mixing framework of Rosenblatt (1956), and the derived *asymptotic distribution is the same for dependent and independent cases*.

    "No effect if ignoring the dependence! Why?"

What about high quantile regression under dependence?

- Chernozhukov (2005, *AoS*):
  - An influential work that considered the possibility of allowing time series data by using the strong mixing framework of Rosenblatt (1956), and the derived *asymptotic distribution is the same for dependent and independent cases*.

    "No effect if ignoring the dependence! Why?"
  - To handle high quantiles, however, besides the mixing condition, an additional condition was required to control the joint probability of nearby tail events; see also Chernozhukov and Fernández-Val (2011). *Such a condition can be interpreted as a negligibility condition on tail dependence.*

    "Dependence is only allowed to exist in non-tail regions!"
- What I contribute:
  1. Propose a new framework based on tail adversarial stability for asymptotic theory of tail dependent time series;
  2. Develop an asymptotic theory for high quantile regression estimators for data with potentially non-negligible tail dependence.

  "Tail dependence indeed affects high quantile regression estimators!"

Tail Adversarial Stability

A New Framework for Asymptotic Theory Allowing Tail Dependence

Background and Motivation

- A fundamental problem in statistics is to develop limit theorems:
  - important for understanding parameter estimators;
  - useful for guiding statistical inference procedures.

- The development of such limit theorems, however, often requires some fundamental assumptions or beliefs about the observed data.

  "Assumption of Independence: most commonly used but not suitable for time series data"

- Rosenblatt (1956) proposed the influential strong mixing condition:
  - coordinates well with a big blocks small blocks argument and leads to the development of various limit theorems;
  - involves a supremum over two sigma algebras and is in general difficult to calculate (Wu, 2005);
  - does not seem to be ideal for high quantile regression as Chernozhukov (2005, *AoS*) imposed an additional negligibility condition to control the degree of tail dependence.

- Wu (2005) proposed an influential alternative based on the functional dependence measure:
  - coordinates well with the projection operator of Hannan (1979) and has been successful in obtaining various limit theorems;
  - easy to calculate, especially for linear processes;
  - tends to summarize the dependence over the whole support and may not be suitable for investigating tail dependence.

"Requires significant innovations over the correlation"

- I propose a new framework:
  - convenient and mathematically rigorous for developing limit theorems of tail dependent time series;
  - based on a previously undescribed tail adversarial stability notion;
  - coordinates well with an m-dependent martingale approximation scheme that leads to desired limit theorems;
  - can be easily calculated for the moving maximum process (Hall, Peng and Yao, 2002).

- Suppose we observe a stationary time series $Y_1, \dots, Y_n$ according to

$$Y_i = G(\mathcal{F}_i), \quad \mathcal{F}_i = (\dots, \epsilon_{i-1}, \epsilon_i),$$

  - G is a measurable function representing the physical system;
  - ϵ_j, $j \in \mathbb{Z}$, are iid innovations (system inputs);
  - covers a wide range of processes (Wiener, 1958 and Wu, 2005).

- Denote the distribution function by F, then for small choices of a, $Y_i > F^{-1}(1-a)$ represents a tail event.

- Let

$$Y_i^\star = G(\mathcal{F}_i^\star), \quad \mathcal{F}_i^\star = (\dots, \epsilon_{-1}, \epsilon_0^\star, \epsilon_1, \dots, \epsilon_i),$$

  we consider

$$\theta_\alpha(i) = \sup_{a \in (0,\alpha]} \mathrm{pr}\{Y_i^\star \le F^{-1}(1-a) \mid Y_i > F^{-1}(1-a)\}.$$

"Measures the tail adversarial effect of current innovation on Y_i"

  - Example: If Y_i does not depend on ϵ_0, then $\theta_\alpha(i) = 0$.

- Recall that

$$Y_i = G(\mathcal{F}_i), \quad \mathcal{F}_i = (\ldots, \epsilon_{-1}, \epsilon_0, \epsilon_1, \ldots, \epsilon_i);$$
$$Y_i^\star = G(\mathcal{F}_i^\star), \quad \mathcal{F}_i^\star = (\ldots, \epsilon_{-1}, \epsilon_0^\star, \epsilon_1, \ldots, \epsilon_i),$$

and $$\theta_\alpha(i) = \sup_{a \in (0,\alpha]} \mathrm{pr}\{Y_i^\star \leq F^{-1}(1-a) \mid Y_i > F^{-1}(1-a)\}.$$

"Measures the tail adversarial effect of current innovation on Y_i"

- For $q \geq 1$, define

$$\Theta_{\alpha,q}(0) = \sum_{i=0}^{\infty} \{\theta_\alpha(i)\}^{1/q}, \quad k \geq 0,$$

that *measures the cumulative tail adversarial effect on future observations*, then a process is said to be tail adversarial q-stable or $(Y_i) \in \mathrm{TAS}_q$ if

$$\lim_{\alpha \downarrow 0} \Theta_{\alpha,q}(0) < \infty.$$

"Finite cumulative tail adversarial effect on future observations"

- Example: Moving maximum process (Hall, Peng and Yao, 2002):

$$Y_i = \max_{0 \le l < \infty} a_l \epsilon_{i-l}, \quad i = 1, \ldots, n,$$

  - ϵ_j: independent Fréchet with $F_\epsilon(z) = \exp(-z^{-\gamma})$, $\gamma > 0$;
  - well defined if $\sum_{l=0}^{\infty} a_l^{\gamma} < \infty$;
  - *proven to be dense in the class of stationary processes whose finite-dimensional distributions are extreme-value of a given type.*

- By elementary calculation, one can show that for α small enough,

$$\theta_\alpha(i) \le \frac{2a_i^{\gamma}}{\sum_{l=0}^{\infty} a_l^{\gamma}}, \quad \Theta_{\alpha,q}(0) \le \frac{2^{1/q} \sum_{i=0}^{\infty} a_i^{\gamma/q}}{(\sum_{l=0}^{\infty} a_l^{\gamma})^{1/q}},$$

and thus $(Y_i) \in \mathrm{TAS}_q$ if

$$\sum_{i=0}^{\infty} a_i^{\gamma/q} < \infty.$$

"Calculating the strong mixing coefficient can be nontrivial and may possibly lead to stronger conditions."

High Quantile Regression under Tail Dependence

Taking Advantage of the Newly Proposed Framework

- Consider the high quantile regression estimator:

$$\hat{\boldsymbol{\beta}}_n = \operatorname*{argmin}_{\boldsymbol{\eta} \in \mathbb{R}^p} \sum_{i=1}^{n} \phi_{1-\alpha_n}(y_{i,n} - \boldsymbol{x}_{i,n}^\top \boldsymbol{\eta}),$$

  - $\phi_{1-\alpha_n}(y) = (1-\alpha_n)y^+ + \alpha_n(-y)^+$ is the check function;
  - $1-\alpha_n \to 1$ as $n \to \infty$ to capture the tail phenomena;
  - allows the observed data to be a triangular array.

Theorem (Zhang, 2021)

Assume certain regularity conditions including the TAS_q, we have

$$\hat{\boldsymbol{\beta}}_n - \boldsymbol{\beta}_n = O_p(\tau_n^{-1}).$$

"τ_n: explicit expression available, depend on n, α_n and tail behavior"

- Examples:

	τ_n
Uniform	$n^{1/2}\alpha_n^{-1/2}$
Normal	$(n\alpha_n)^{1/2}\{\log(\alpha_n^{-1})\}^{1/2}$
Exponential	$(n\alpha_n)^{1/2}$
Pareto	$(n\alpha_n^{1+2\lambda})^{1/2}$

"Convergence rate is no longer the universal $n^{1/2}$, and in certain cases can even exceed the parametric rate"

Theorem (Zhang, 2021)

Assume certain regularity conditions including the TAS_q, we have

$$\tau_n(\hat{\boldsymbol{\beta}}_n - \boldsymbol{\beta}_n) \to_d N(0, \boldsymbol{\Gamma}),$$

where

$$\boldsymbol{\Gamma} = \sum_{k \in \mathbb{Z}} \rho_k \boldsymbol{\Upsilon}_0^{-1} \boldsymbol{\Upsilon}_k \boldsymbol{\Upsilon}_0^{-1}.$$

- $\boldsymbol{\Upsilon}_k$: *limit of* $n^{-1} \sum_{i=1}^{n-|k|} \boldsymbol{x}_{i,n} \boldsymbol{x}_{i+|k|,n}^\top$ *as* $n \to \infty$;
- ρ_k: *standardized lag-k tail dependence index of Zhang (2005).*

- Example: For independent data, the above central limit theorem reduces to

$$\tau_n(\hat{\boldsymbol{\beta}}_n - \boldsymbol{\beta}_n) \to_d N(0, \boldsymbol{\Upsilon}_0^{-1}).$$

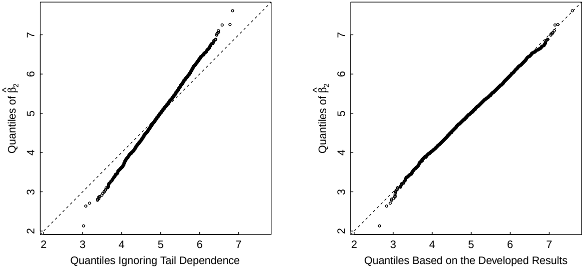
"Ignoring the dependence can lead to misleading p-values"

Simulation Studies and Data Analysis

Numerical Experiments

A Simulation Study

- Consider:
  - Trend analysis design: $\boldsymbol{x}_{i,n} = (1, t_{i,n})^\top$, $t_{i,n} = i/n$;
  - Moving maximum dependence: $U_{i,n} = \max(\epsilon_i, \epsilon_{i-1}) - a_n$;
  - Focus on the distribution of $\hat{\beta}_2$, the coefficient of $t_{i,n}$.



"Ignoring the dependence can lead to misleading p-values!"

	0.1	0.05	0.01
Ignoring the dependence	0.224	0.150	0.059
Developed results	0.087	0.043	0.007

Application to a Temperature Data

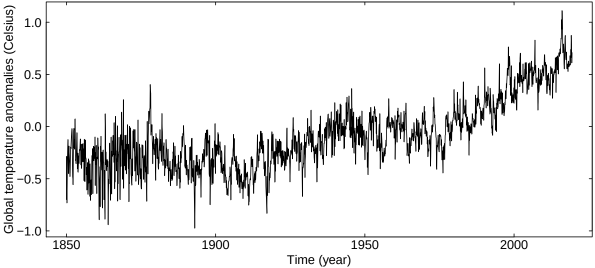


Figure: Monthly global temperature anomalies in Celsius from 1850 to 2019

- Temperature is an important factor in economics (Dell, Jones and Olken, 2012).
- Rust (2003) and Wu and Zhao (2007, *JRSSB*):
  - *quadratic trend for the mean function.*
- Q: Is the quadratic trend still enough for high quantiles or if a higher order polynomial (e.g. cubic) is needed?

Application to a Temperature Data (Continued)

Testing if a cubic trend can be reduced to a quadratic one

	95% quantile		*99% quantile*	
Tail dependence	Adjusted	Ignored	Adjusted	Ignored
Cubic Coefficient Estimate	1.211	1.211	3.421	3.421
p-value	0.270	0.008	0.002	0.000

Table: High quantile regression estimators for the cubic coefficient and their associated p-values for testing a zero null hypothesis against a two-sided alternative.

- Findings:
  - Ignoring the dependence ⇒ underestimated uncertainty ⇒ distortedly narrower confidence intervals

    "can be a serious problem in risk management"
  - the quadratic form continues to be adequate for the 95% quantile, while a higher order polynomial is needed for the 99% quantile.

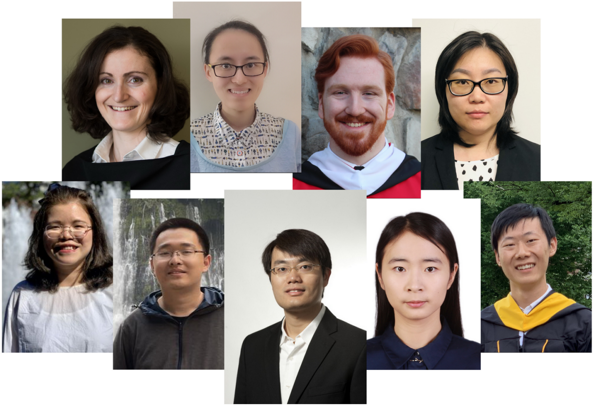
    "the trend in 99% quantile can be more complicated"

Zhang, T. (2021).
High-quantile regression for tail-dependent time series.
Biometrika, **108**, 113–126.

To summarize, in this work I make two major contributions.

1. I propose a new framework:
   - convenient and mathematically rigorous foundation for developing limit theorems of tail dependent time series;
   - based on a previously undescribed tail adversarial stability notion;
   - coordinates well with an m-dependent martingale approximation scheme that leads to desired limit theorems;
   - can be easily calculated for the moving maximum process (Hall, Peng and Yao, 2002).
2. Taking advantage of the newly proposed framework:
   - I consider the problem of high quantile regression for time series data;
   - unlike the result of Chernozhukov (2005, AoS), dependence is allowed in both tail and non-tail regions;
   - the associated convergence rate and a central limit theorem are established for a general class of tail dependent processes.

Questions?



Thank You!

Zhang's Statistics Lab: tingzhang@uga.edu