Finite sample valid probabilistic inference on quantile regression

ISIPTA 2023

Leonardo Cella: www.leocella.com

WAKE FOREST UNIVERSITY

Introduction

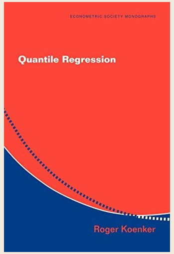
Setup: Data $Z^n = (Z_1, \ldots, Z_n)$, $Z_i = (X_i, Y_i)$, are iid with distribution P.

Quantile Regression:

- Let $Q_x(\tau)$ denote the **τ**-th quantile of Y given X=x
- $Q_x(\tau) = x^\top \theta$
- **Goal**: Make inferences on θ → distribution free; → valid

Traditional methods

- Inferences through <u>confidence regions</u> →



- Notion of validity is familiar → coverage guarantees:

$$\sup_P P^n\{C_\alpha(Z^n) \not\ni \theta(P)\} \leq \alpha, \quad \alpha \in [0,1].$$

- Traditional methods usually achieve this <u>asymptotically</u>

Can we do more?

Probabilistic inference

- Assign *degrees of belief* to general assertions about θ

- **Validity**: Control *the assignment of **high** degrees of belief to **false** assertions*

$$\sup_{P:\theta(P)\notin A} P^n\{\Pi_{Z^n}(A) > 1-\alpha\} \leq \alpha, \quad \alpha \in [0,1].$$

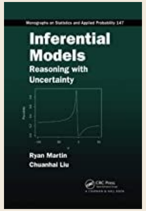
- Bayes comes to mind, but ***False confidence theorem*** says we need imprecision!

Satellite conjunction analysis and the false confidence theorem

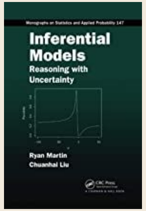
Michael Scott Balch, Ryan Martin and Scott Ferson

Published: 17 July 2019 | https://doi.org/10.1098/rspa.2018.0565

"Degrees of belief that are (additive) probabilities are at risk of being invalid"



Inferential Models



Imprecise probabilistic approach to statistical inference that has validity as its main pilar

Overview

Setup: Data $Z^n = (Z_1, \ldots, Z_n)$ is iid with distribution Pω

Two-step IM construction

1. Choose an appropriate $h : (\mathbb{Z}^n \times \Omega) \to \mathbb{R}$ that determines a partial ordering of candidate values for ω given z^n, e.g.,likelihood ratio:

$$h(z^n, \omega) = L_{z^n}(\omega)/L_{z^n}(\hat{\omega}_{z^n})$$

2. Compute the possibility contour

$$\pi_{z^n}(\omega) = P_\omega^n\{h(Z^n, \omega) \le h(z^n, \omega)\}$$

Degrees of belief

Confidence Regions

Back to quantile regression...

- Data $Z^n = (Z_1, \ldots, Z_n)$ are iid with distribution P
- $Z_i = (X_i, Y_i)$
- $Q_x(\tau) = x^T\theta \rightarrow \tau$-th quantile of Y given X=x
- $\theta = \theta(P)$: **Inferential target**

IM construction

1- Choose an appropriate h that orders candidate values for θ given data

2 - Compute the contour $\pi_{z^n}(\theta) = P^n\{h(Z^n, \theta) \leq h(z^n, \theta)\}$

Theorem:
This construction yields valid degrees of belief for assertions about θ, as well as valid confidence regions for θ.

Problems....

1- Choose h

No model

No likelihood ratio

2 - Compute the contour

$$\pi_{z^n}(\theta) = P^n\{h(Z^n,\theta) \leq h(z^n,\theta)\}$$

P is not known

Possible solution


ELSEVIER

International Journal of Approximate Reasoning

Volume 151, December 2022, Pages 205-224

Direct and approximately valid probabilistic inference on a class of statistical functionals

Leonardo Cella, Ryan Martin

h in terms of relative risk

π approximated by the bootstrap

Due to the bootstrap, validity is just achieved asymptotically!

Basic idea

Choose an h whose distribution is known and independent of unknown quantities

↓

$\pi_{z^n}(\theta) = P^n\{h(Z^n, \theta) \leq h(z^n, \theta)\}$ **can be computed**

But how to obtain such h?

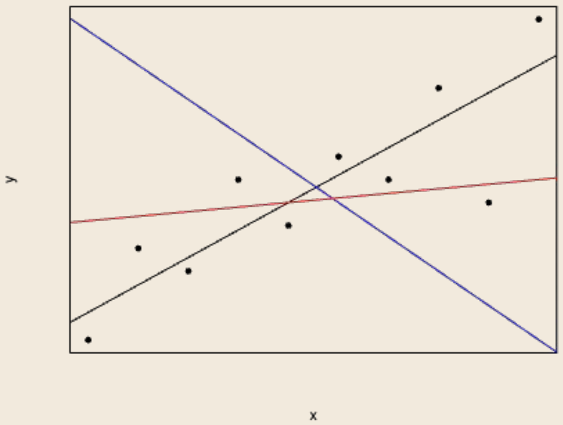
1) Find a function γ of data and θ that is a pivot

2) Set h as the γ's probability mass

Intuitively...

$$\gamma = \sum_{i=1}^{n} I_{(0,\infty)}(Y_i - x_i^\top \theta) \Rightarrow \gamma \sim \mathrm{Bin}(n, 1-\tau) \Rightarrow h = \binom{n}{\gamma}(1-\tau)^{\gamma}\tau^{n-\gamma}$$

But not efficient...



New consideration: Discrete case first

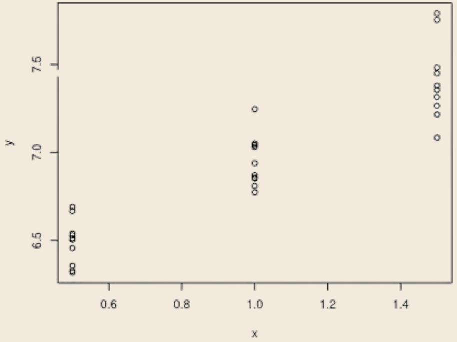
Consider the independent binomials at each level of X separately!

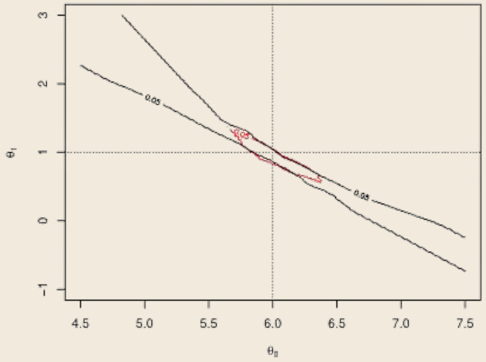
$$h = \prod_{i=1}^{k} \binom{n_i}{\gamma_i} (1-\tau)^{\gamma_i} \tau^{n_i - \gamma_i},$$

where

$$\gamma_i = \sum_{j=1}^{n_i} I_{(0,\infty)}(Y_j - x_i\theta), \quad i = 1, \ldots, k.$$

Example: n=30, τ=0.5





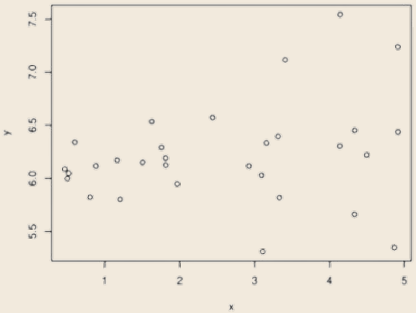
Continuous X

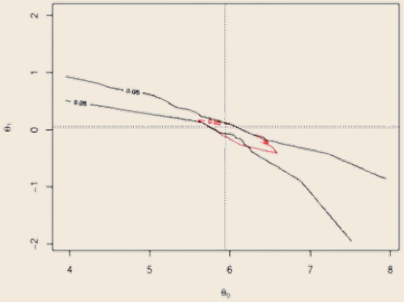
If X is continuous, **there is no replication** of Y for any given X = x

But we do have replications of Y in neighborhoods of X!

- Form k neighborhoods of X
- Consider each one of the k independent binomials separately
- h: product of their probability masses

Example: n=30, τ=0.3, k=2





This was fast, but….

- More details
- More examples
- Comparisons with other methods
- Open questions

Finite sample valid probabilistic inference on quantile regression

Leonardo Cella CELLAL@WFU.EDU
Department of Statistical Sciences, Wake Forest University, USA