Econometria I
Regressão Quantílica

Victor Oliveira
Núcleo de Economia Internacional e Desenvolvimento Econômico

PPGDE – 2023

Sumário I

Introdução

- Quando falamos de regressão, geralmente nos referimos à regressão média, que descreve como o valor esperado de uma variável de resposta de interesse varia com as variáveis explicativas.
- O que a curva de regressão faz é fornecer um grande resumo das médias das distribuições correspondentes ao conjunto da matriz $\boldsymbol{X}$.
- Poderíamos ir mais longe e calcular várias curvas de regressão diferentes correspondentes aos vários pontos percentuais das distribuições e, assim, obter uma imagem mais completa do conjunto.
- Normalmente, isso não é feito e, portanto, a regressão geralmente fornece um quadro bastante incompleto. Assim como a média fornece uma imagem incompleta de uma única distribuição, a curva de regressão fornece uma imagem correspondentemente incompleta para um conjunto de distribuições.

- Quando falamos de regressão, geralmente nos referimos à regressão média, que descreve como o valor esperado de uma variável de resposta de interesse varia com as variáveis explicativas.
- Às vezes também é útil considerar o efeito das variáveis explicativas em todo o distribuição condicional da variável de interesse.
- A regressão quantílica é uma maneira de conseguirmos isso.
- Introduziremos a regressão quantílica considerando primeiro o que é um quantil e o que se entende por regressão quantil, antes de apresentar os principais propriedades dos quantis e a teoria da estimativa e inferência para regressão quantílica
- Uma característica atraente da regressão quantílica que tem sido repetidamente enfatizada é que ela nos permite olhar para "fatias" da distribuição condicional sem qualquer dependência de suposições distributivas globais.

Quantis

Definição

Suponha que a variável aleatória Y tenha uma função de distribuição cumulativa $F_Y(y) = \mathrm{P}(Y \leq y)$. O τ-ésimo quantil de Y é definido como

$$Q_\tau(Y) = \inf\{y \colon F_Y(y) \geq \tau\} \tag{1}$$

com $0 < \tau < 1$.

- $Q_\tau(Y)$ é uma estatística de ordem, e pode ser obtida pela minimização de uma função perda (linear) assimétrica:

$$\tau \int_{y>\tau} |y-\tau| dF_Y(y) + (1-\tau) \int_{y<\tau} dF_Y(y) \qquad (2)$$

- A contraparte da função perda (linear) assimétrica é dada por

$$\frac{1}{N}\sum_{i=1}^{N} \rho_\tau(y_i - \tau) = \frac{1}{N}\sum_{i=1}^{N} \left[\tau \sum_{n:\, y_n \geq \tau} |y_n - \tau| + \right.$$
$$\left. (1-\theta) \sum_{n:\, y_n < \tau} |y_n - \tau| \right] \qquad (3)$$

Função check

- A função check é uma função de perda que recupera o τ-ésimo quantil da amostra.
- A função check dá pesos assimétricos ao erro dependendo do quantil e do sinal geral do erro.
- O τ-ésimo quantil amostral de Y resolve

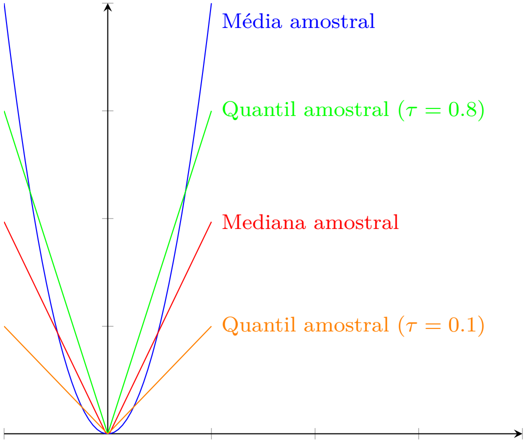
$$\min_{a} \sum_{i=1}^{n} \rho_\tau(y_i - a) \tag{4}$$

- Por exemplo, se você deseja o quantil $\tau = 0.1$, isso significa que 90% dos erros devem ser positivos e 10% negativos.
- Para encontrar o LAD ótimo enquanto essa afirmação é verdadeira, os pesos devem ser adicionados aos erros.
- No caso do 10º quantil, um peso de 0,9 é adicionado aos pesos negativos enquanto um peso de 0,1 é adicionado aos positivos.

Quantis

- Da definição de um quantil podemos ver que $Q_{0.5}(Y)$ é a mediana, também referida como o segundo quartil, enquanto $Q_{0.25}(Y)$ é o primeiro quartil ou 25º percentil e $Q_{0.75}(Y)$ é o terceiro quartil ou 75º percentil.
- Quantis e percentis são essencialmente a mesma coisa, exceto que os primeiros se referem a proporções enquanto os últimos a porcentagens.
- A função quantílica $Q_\tau(Y)$ é uma função não decrescente de τ, ou seja, $Q_{\tau_1}(Y) \leq Q_{\tau_2}(Y)$ para $\tau_1 < \tau_2$.
- Em um arcabouço de regressão, estamos realmente interessados no τ-ésimo quantil condicional.

Visão Geométrica da Otimização

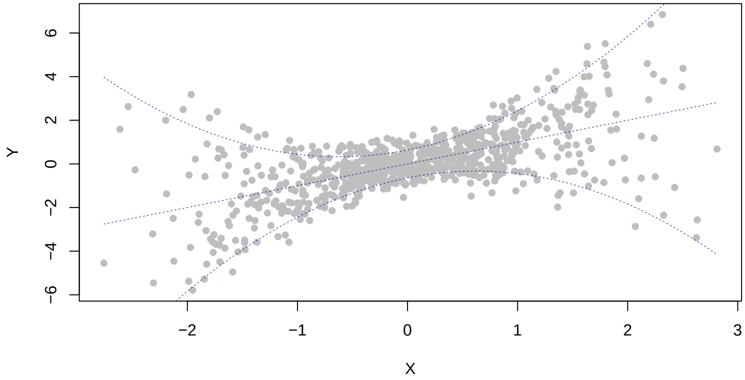


Propriedades

- Uma propriedade das funções de regressão quantílica é que elas são monotonicamente crescentes em τ.
- Isso significa que funções quantílicas para diferentes quantis não podem se cruzar.
- No entanto, uma propriedade das funções lineares $\boldsymbol{X}'\boldsymbol{\beta}$ com inclinações diferentes é que elas necessariamente se cruzarão se o suporte para $\boldsymbol{X}$ for suficientemente grande.
- Este é um problema potencial em aplicações, pois usos práticos de funções quantílicas estimadas podem exigir monotonicidade em τ (por exemplo, se forem invertidas para obter uma função de distribuição condicional).
- Isso é apenas um problema em aplicações práticas se as funções quantílicas estimadas realmente se cruzarem. Se não o fizerem, esse problema pode ser ignorado. No entanto, quando as funções de regressão quantílicas estimadas se cruzam, pode ser prudente resolver o problema.

Exemplo

Figura 1: Quatiles Crossing



- Assim, dada uma amostra $\{Y_1, \ldots, Y_n\}$ de uma distribuição F, pode ser mostrado que
  1. Média amostral: $\bar{Y} = \arg\min_a \sum_{i=1}^{n} (Y_i - a)^2$
  2. Mediana amostral: $\widehat{Q}_Y(0.5) = \arg\min_a \sum_{i=1}^{n} |Y_i - a|^2$
  3. τ-ésimo quantil amostral: $\bar{Y} = \arg\min_a \sum_{i=1}^{n} \rho_\tau(Y_i - a)$

Definição

Suponha que tenhamos a variável aleatória Y e que $\boldsymbol{X}$ é um preditor p-dimensional. Seja $F_Y(y|\boldsymbol{X}) = \mathrm{P}(Y \leq y|\boldsymbol{X})$ denota a FDA condicional de Y dado $\boldsymbol{X}$. Então, o τ-ésimo quantil condicional de Y é definido como

$$Q_\tau(Y|\boldsymbol{X}) = \inf\{y \colon F_Y(y|\boldsymbol{X}) \geq \tau\} \tag{5}$$

com $0 < \tau < 1$.

- Isto nos permite considerar o modelo da forma

$$Q_\tau(Y|\boldsymbol{X}) = \boldsymbol{X}'\boldsymbol{\beta}(\tau) \tag{6}$$

com média $\mathbb{E}[Y|\boldsymbol{X}] = \boldsymbol{X}'\boldsymbol{\beta}$.

- Podemos interpretar $\boldsymbol{\beta}(\tau)$ como a mudança marginal no τ-ésimo quantil devido à mudança marginal em $\boldsymbol{X}$.
- Observe que isso é para um valor específico de τ e que diferentes quantis podem ter coeficientes que diferem entre si em magnitude, sinal ou ambos.
- Observe também a propriedade de monotonicidade dos quantis condicionais: $Q_\tau(Y|\boldsymbol{X})$ é uma função não decrescente de τ para qualquer $\boldsymbol{X}$ dado.

- Devido à função perda, $\widehat{\beta}(\tau)$ é mais robusto para *outliers* do que o estimador de MQO.
- A regressão quantílica não é a mesma que as regressões baseadas em amostras divididas porque cada regressão quantílica utiliza todos os dados de amostra (com diferentes pesos). Assim, regressão quantílica também evita o problema de seleção de amostra decorrente da divisão da amostra.

Propriedades

- Considere $\widehat{\beta}(\tau; y, \boldsymbol{X})$ o estimador para a regressão do τ-ésimo quantil baseado nas observações $(y, \boldsymbol{X})$ e seja A uma matriz não singular $p \times p$, $\gamma \in \mathbb{R}^p$, e $a > 0$ uma constante. Então, para qualquer $\tau \in [0, 1]$,
  1. $\widehat{\beta}(\tau; ay, \boldsymbol{X}) = a\widehat{\beta}(\tau; y, \boldsymbol{X})$ e $\widehat{\beta}(\tau; -ay, \boldsymbol{X}) = -a\widehat{\beta}(1-\tau; ay, \boldsymbol{X})$: escala de equivariância
  2. $\widehat{\beta}(\tau; y + \boldsymbol{X}\gamma, \boldsymbol{X}) = \widehat{\beta}(\tau; y, \boldsymbol{X}) + \gamma$: regressão shift
  3. $\widehat{\beta}(\tau; y, \boldsymbol{X}A) = A^{-1}\widehat{\beta}(\tau; y, \boldsymbol{X})$: reparametrização do design

- Além disso, as funções quantílicas condicionais são equivalentes às transformações monótonas. Suponha que $h(\cdot)$ seja uma função crescente em $\mathbb{R}$. Então, para qualquer variável Y,

$$Q_\tau(h(Y|\boldsymbol{X})) = h(Q_\tau(Y|\boldsymbol{X})) \tag{7}$$

- Ou seja, os quantis da variável aleatória transformada $h(Y)$ são simplesmente os quantis transformados na escala original. Isso é útil, por exemplo, quando transformamos em log a resposta.
- Observe que $Q_\tau(\log(Y|\boldsymbol{X})) = \log(Q_\tau(Y|\boldsymbol{X}))$, mas $\mathbb{E}(\log(Y)|\boldsymbol{X}) \neq \log(\mathbb{E}(Y|\boldsymbol{X}))$.

Estimação

- Suponha que observamos realizações $\{y_i\}$ com valores correspondentes das variáveis explicativas $\boldsymbol{X}$.
- Na regressão média usando estimativa de mínimos quadrados ordinários (OLS), estimamos os coeficientes de regressão minimizando a soma dos quadrados. O caso mais simples é o de um modelo somente com intercepto, em que $\mathbb{E}[Y] = \mu_Y = \arg\min_a \mathbb{E}(Y-a)^2$. A média amostral $\min_a \sum_{i=1}^n (y_i - a)^2$. Com outras variáveis, do estimador de MQO $\min \sum_{i=1}^n (y_i - \boldsymbol{X}'\boldsymbol{\beta})^2$.
- No caso de regressão quantílica, usamos o método de desvios absolutos mínimos (LAD).
- Para a mediana condicional resolvemos

$$\min_{\beta} \sum_{i=1}^{n} |y_i - \boldsymbol{X}'\boldsymbol{\beta}| \tag{8}$$

- Considere a regressão quantílica para o quantil $0 < \tau < 1$. O τ-ésimo quantil de Y é dado por

$$Q_\tau(Y) = \arg\min_a \mathbb{E}[\rho_\tau(Y - a)] \tag{9}$$

em que $\rho_\tau(u) = u\tau - I(u < 0)$ é a função perda.

Propriedades Estatísticas

- Se assumirmos que $Q_\tau(Y|\boldsymbol{X}) = \boldsymbol{X}'\boldsymbol{\beta}(\tau)$, então

$$\widehat{\boldsymbol{\beta}}(\tau) = \arg\min_{\beta} \sum_{=1}^{n} \rho_\tau(y - \boldsymbol{X}'\boldsymbol{\beta}) \tag{10}$$

- As estimativas do coeficiente de regressão quantílica são consistentes e seguem uma distribuição normal assintótica. O estimador de coeficiente em um modelo de regressão quantílica linear é dado por

$$\widehat{\boldsymbol{\beta}} = \arg\min_{\boldsymbol{b}\in\mathbb{R}^p} \sum_{i=1}^{n} \rho_\tau(y_i - \boldsymbol{X}'\beta) \tag{11}$$

- Os coeficientes estimados são consistentes sob as seguintes condições de regularidade

  CR1 As funções de distribuição de Y dado $\boldsymbol{X}$, $F(\cdot)$, são absolutamente contínuas com densidade contínua $f(\cdot)$ que são uniformemente limitadas entre 0 e ∞ em $\boldsymbol{\xi} = Q_\tau(Y|\boldsymbol{X})$.

  CR2 Existem matrizes D_0 e D_1 positiva definidas tal que

  1. $\lim_{n\to\infty} \mathbb{E}[\boldsymbol{X}'\boldsymbol{X}] = D_0$
  2. $\lim_{n\to\infty} \mathbb{E}[\boldsymbol{f}(\boldsymbol{\xi}(\tau))\boldsymbol{X}'\boldsymbol{X}] = D_1(\tau)$
  3. $\max \|\boldsymbol{X}\| = o\left(n^{1/2}\right)$

- Sob as condições CR1 e CR2(1), $\widehat{\boldsymbol{\beta}}(\tau) \xrightarrow{P} \widehat{\beta}(\tau)$.

- Sob as condições CR1 e CR2,

$$\sqrt{n}\left(\widehat{\boldsymbol{\beta}}(\tau) - \widehat{\beta}(\tau)\right) \xrightarrow{D} \mathcal{N}\left(0, \tau(1-\tau)D_1^{-1}D_0D_1^{-1}\right) \tag{12}$$

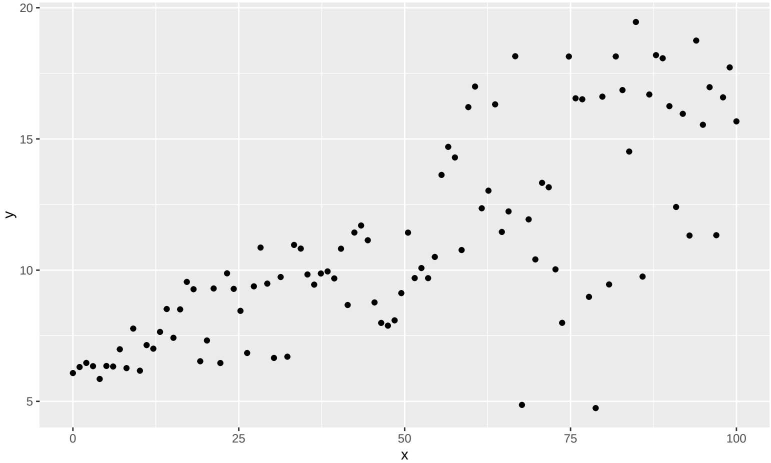
- Para um modelo homocedástico, isto é, $\boldsymbol{f}(\boldsymbol{\xi}(\tau)) = \boldsymbol{f}_\varepsilon(0)$, o resultado acima é simplificado para

$$\sqrt{n}\left(\widehat{\boldsymbol{\beta}}(\tau) - \widehat{\beta}(\tau)\right) \xrightarrow{D} \mathcal{N}\left(0, \frac{\tau(1-\tau)}{\boldsymbol{f}_\varepsilon(0)}D_0\right) \tag{13}$$

em que $\boldsymbol{f}$ é estimado de forma não paramétrica.

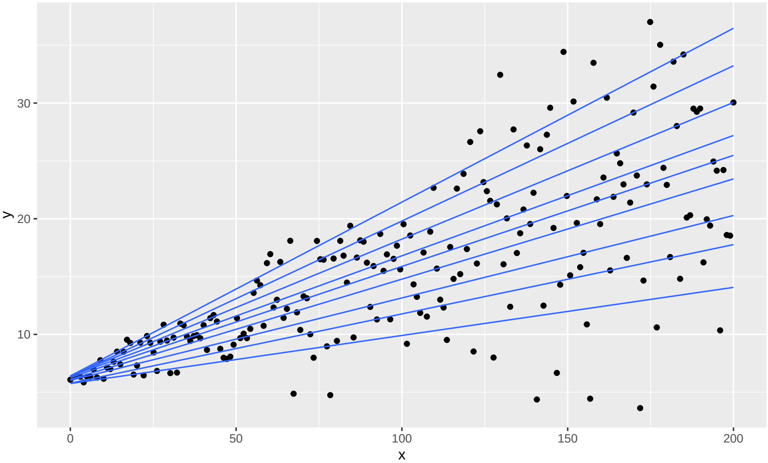
Exemplo

Figura 2: Distribuição Conjunta dos Dados



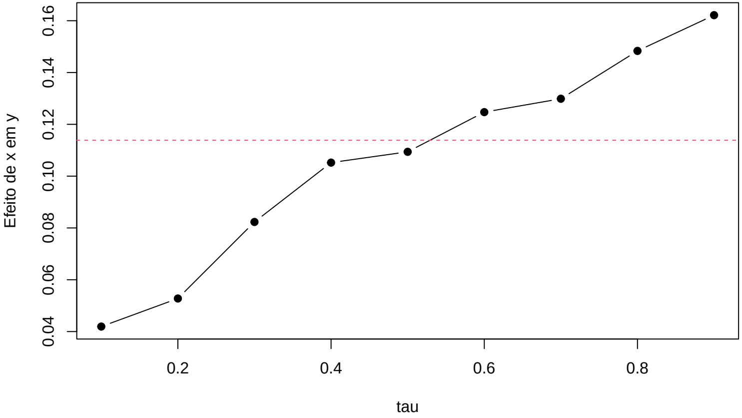
Exemplo

Figura 3: Regressão Quantílica



Exemplo

Figura 4: Sequência de β's



Econometria I
Regressão Quantílica

Victor Oliveira
Núcleo de Economia Internacional e Desenvolvimento Econômico

PPGDE – 2023