

Regressão Quantílica

Quando utilizar!?

*Luiz Fernando Palin Droubi**

Carlos Augusto Zilli†

Murilo Damian Ribeiro‡

Norberto Hochheim§

07/01/2020

Resumo

A NBR 14.653-02 ([ABNT, 2011](#)) recomenda que, na Engenharia de Avaliações de imóveis urbanos, para o tratamento dos dados seja utilizada metodologia científica, mesmo no tratamento de dados por fatores, o que usualmente é feito através do método da regressão linear clássica ou ordinária, ainda que a norma também cite outros métodos, como a regressão espacial, a análise envoltória de dados e as redes neurais artificiais. No entanto, através destes métodos, o que se obtém são coeficientes ou fatores **médios** da contribuição de uma característica do imóvel na formação do valor final. Ocorre que a contribuição de uma determinada característica para a formação do valor final dos imóveis pode ser diferente para os diferentes quantis da distribuição de probabilidades. É possível até que uma determinada característica que se mostre insignificante no método da regressão linear seja significativa na regressão quantílica, já que na regressão linear o que se estima é se, **em média**, uma determinada característica tem influência na formação do valor total de um imóvel. Ocorre que uma determinada característica pode influenciar positivamente o preço de venda dos imóveis de maior valor e negativamente o preço de venda dos imóveis de menor valor (ou *vice versa*), sendo que, **em média**, o seu efeito é nulo, o que no entanto não quer dizer que aquela variável não tenha qualquer influência na formação de preço dos imóveis do mercado em análise. Em suma, esta diferente contribuição das características no valor final dos imóveis, atualmente, é ignorada (utiliza-se apenas o valor médio), fazendo com que eventuais diferenças nos efeitos das características em imóveis de valores diferentes sejam negligenciadas. A regressão quantílica é um método que permite estimar a real influência de cada característica ao longo de toda a distribuição de probabilidades dos imóveis de um mercado, o que pode se demonstrar útil na avaliação de imóveis urbanos em determinados mercados o que atualmente pode passar despercebido aos olhos do avaliador acostumado com os métodos estatísticos clássicos.

*SPU/SC, lfpdroubi@gmail.com

†IFSC, carloszilli@gmail.com

‡UFSC, murilodamianr@gmail.com

§UFSC, hochheim@gmail.com

1 Introdução

What the regression curve does is give a grand summary for the averages of the distributions corresponding to the set of x 's. We could go further and compute several different regression curves corresponding to the various percentage points of the distributions and thus get a more complete picture of the set. Ordinarily this is not done, and so regression often gives a rather incomplete picture. Just as the mean gives an incomplete picture of a single distribution, so the regression curve gives a correspondingly incomplete picture for a set of distributions. — [Mosteller e Tukey \(1977\)](#) (apud [KOENKER, 2000](#), p. 19)

Em muitas áreas da Ciência, como na Ecologia, é possível que um pesquisador, ao analisar dados como os da figura 1 com a utilização do consagrado método da regressão linear, chegue à conclusão de que não há qualquer correlação entre duas variáveis, já que, pelo estudo das variáveis, *em média*, o efeito do regressor encontrado é nulo ($\beta = 0$), tal como mostra a reta de regressão linear (tracejada) na figura.

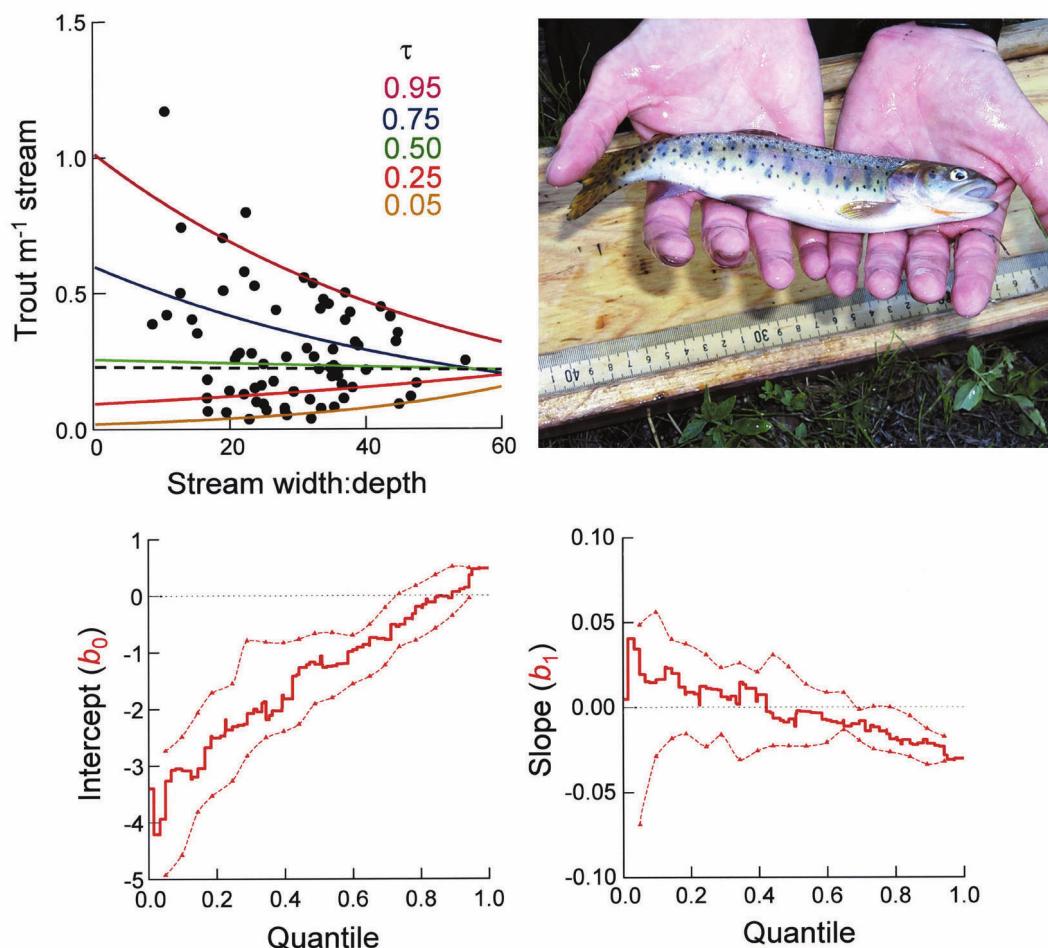


Figura 1: Mudanças na densidade de trutas (y) em função do quociente largura/profundidade de um canal (X). Fonte: [Cade e Noon \(2003, p. 413\)](#).

No entanto, como bem observaram [Cade e Noon \(2003, p. 412-413\)](#), se os autores

deste estudo tivessem se baseado *apenas* no método da regressão linear, eles teriam fatalmente chegado a esta conclusão errônea. Na realidade, porém, com o auxílio da regressão quantílica, mostrou-se que a relação existe, porém pode ser percebida apenas nos quantis superiores dos dados. Fisicamente, o que deve ser deduzido é que, com o aumento do quociente largura/profundidade de um canal, limita-se a densidade de trutas no canal. Esta limitação não é percebida nos quantis inferiores dos dados, de maneira que, *em média*, o seu efeito é nulo, mas claramente o efeito é considerável nos quantis superiores dos dados. Isto ocorre porque efeitos de fatores limitantes não são bem representados pela média das distribuições de probabilidades, onde a presença de muitos outros fatores limitantes não-medidos podem estar presentes (CADE; NOON, 2003, p. 413).

O mesmo comportamento pode ou poderia ser encontrado na Engenharia de Avaliações. Imagine-se que ao analisar um mercado de lotes urbanos o Engenheiro de Avaliações se depare com os dados mostrados na figura 2. Ao analisar os dados através da regressão linear à média (reta tracejada), o Engenheiro poderia, erroneamente concluir que a área não teria nenhum efeito sobre o valor unitário dos lotes, quando na realidade o que ocorre é que o efeito da variável área é o de aumentar suavemente o valor unitário dos lotes de menor valor, e diminuir de uma forma um pouco mais agressiva o valor unitário dos lotes dos quantis superiores.

Em suma, globalmente, o efeito médio é zero, o que não significa que a variável não possua qualquer significância na formação de preços dos imóveis.

A regressão quantílica permite que a influência de uma característica qualquer de um imóvel tenha efeitos diferentes para diferentes faixas de valores de imóveis.

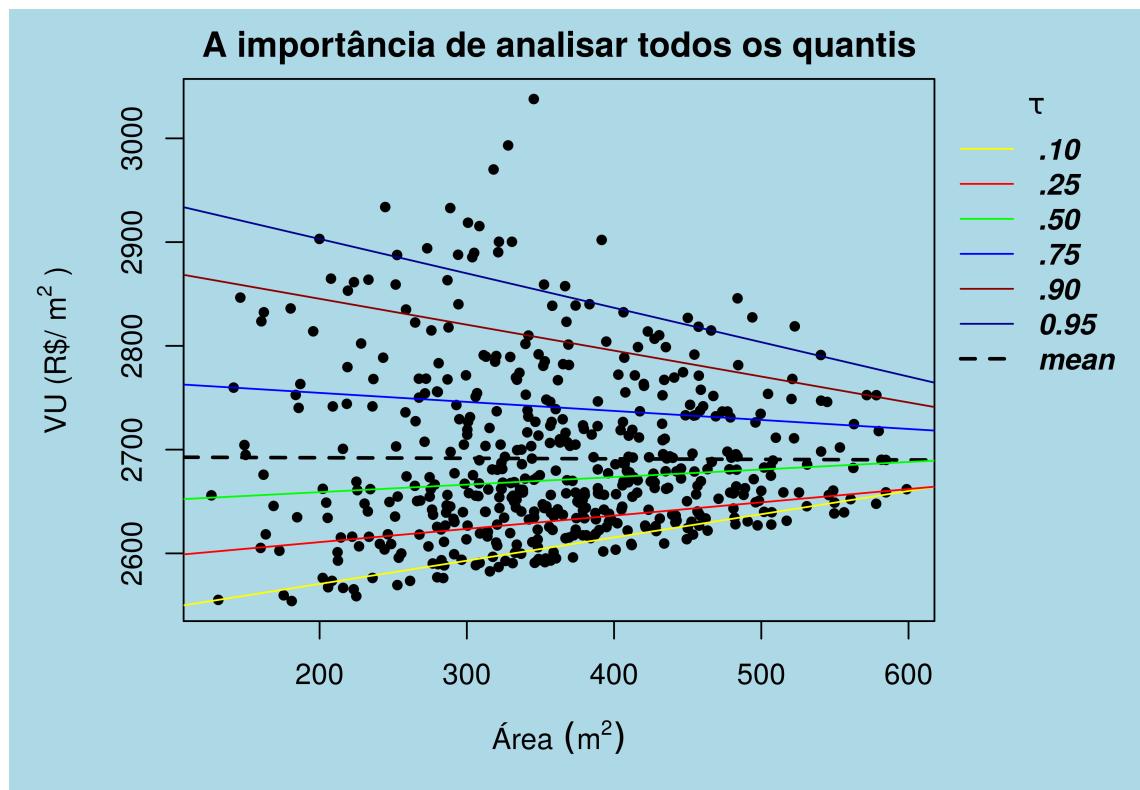


Figura 2: Dados gerados randomicamente para ilustrar a importância da regressão quantílica na Engenharia de Avaliações. Fonte: do Autor.

Joachim Zietz, Emily Norman Zietz e Sirmans (2008) mostram que os conflitos a respeito da contribuição de uma determinada característica na formação dos preços de venda de imóveis residenciais podem ser esclarecidos através da regressão quantílica. Diferentes valores para os coeficientes de regressão linear para algumas características podem ser encontrados ao longo da distribuição de preços de venda de imóveis. Ou seja, algumas características dos imóveis residenciais podem ser mais valorizados por compradores de imóveis de mais alto valor do que por compradores de imóveis de menor valor.

Segundo Joachim Zietz, Emily Norman Zietz e Sirmans (2008), variáveis como área construída, área do lote e número de banheiros tem um impacto maior nos imóveis de maior valor de venda, enquanto outras variáveis parecem ter um comportamento constante para todos os preços de venda de imóveis, como garagens e distância ao centro, entre outras.

2 Revisão Bibliográfica

2.1 Breve Histórico

Segundo Koenker (2000, p. 347), o gráfico mais importante da história da estatística é o gráfico de Galton, reproduzido na figura 3.

O gráfico ilustra o fenômeno, descoberto por Galton, da regressão à média, cuja importância até hoje se faz presente em diversos estudos científicos. Em estudos clínicos para a aferição do real efeito de um novo medicamento, por exemplo, há necessidade de se estabelecerem dois grupos de pacientes (chamados de grupos de controle e de grupo de tratamento) para isolar os efeitos do tratamento pesquisado do efeito da regressão à média (ou reversão à média) (ver JAMES, 1973). Nestes estudos, apenas as pessoas do grupo de tratamento realmente são tratadas com o medicamento em teste. Desta maneira, a diferença das médias entre os grupos é isolada da variação biológica natural e da variação devido aos erros de aferição, que estão sujeitos à regressão à média.

Segundo Koenker (2000, p. 349), a característica essencial da regressão linear clássica, derivada deste gráfico, é que o efeito do covariante na variável resposta é inteiramente capturado pela expressão abaixo, uma simples “mudança de local”, enquanto a aleatoriedade remanescente de Y dado X é modelada por um termo de erro aditivo independente de X .

$$\mathbb{E}(Y|X = x) = x'\beta$$

Talvez prevendo que o seu método fizesse os seus colegas estatísticos se aterem apenas ao estudo das médias, Galton (1889, p. 62) (*apud KOENKER, 2000, p. 350*) repreendeu os seus colegas que:

limited their inquiries to Averages, and do not seem to revel in more comprehensive views. Their souls seem as dull to the charm of variety as that of a native of one of our flat English counties, whose retrospect of Switzerland was

Plate X.

DIAGRAM BASED ON TABLE I.
 (all female heights are multiplied by 108)

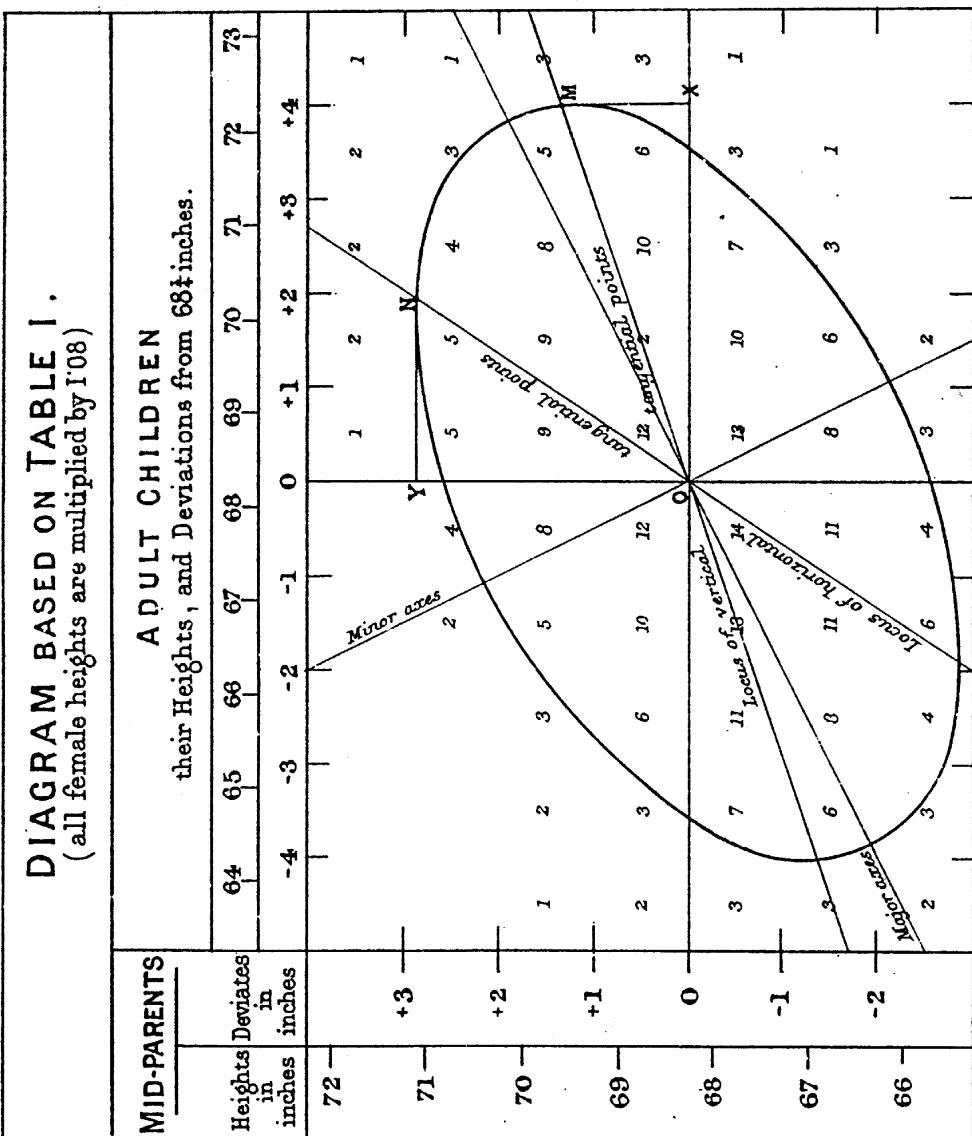


Figura 3: Regressão à média (Galton, 1885). Fonte: Koenker (2000, p. 348)

that, if the mountains could be thrown into its lakes, two nuisances would be got rid of at once.

Ironicamente, contudo, apesar da repreensão de Galton, muito provavelmente foi o seu método o principal responsável por restringir o escopo das investigações estatísticas à comparação de médias por décadas (KOENKER, 2000, p. 350).

Muito anterior ao descobrimento por Galton (1889) do fenômeno da regressão à média e ainda anterior ao descobrimento do método dos mínimos quadrados por Legendre (1805)¹, Boscovich propôs, em 1760, o seguinte problema, em busca de estimar se a hipótese elipsoidal da forma da terra, prevista na obra-prima de Newton (1687), estaria correta, em função de cinco observações do comprimento de um grau de latitude obtidas em diferentes longitudes (KOENKER, 2000, p. 353; PORTNOY; KOENKER, 1997, p. 281; STIGLER, 1986, p. 40):

Encontrar $\hat{\alpha}$ e $\hat{\beta}$ tais que:

$$y_i = \hat{\alpha} + \hat{\beta}x_i + \hat{u}_i$$

com $\sum \hat{u}_i = 0$ e $\sum |\hat{u}_i| = \min!$.

Boscovich (1770) propôs uma solução através de um algoritmo geométrico. Alguns anos depois, em 1789, Laplace (1793) resolveu o problema matematicamente, nomeando o método de “Methode de Situation” no que pode ser considerada a primeira análise de regressão (PORTNOY; KOENKER, 1997, p. 281).

Posteriormente, com a chegada do *méthode des moindres carrés*, ou seja, do método dos mínimos quadrados ordinários, o método dos mínimos erros absolutos de Boscovich/Laplace ficou em segundo plano. Posteriormente, Edgeworth (1887) argumentou que, quando os erros não seguem a lei de Gauss, a regressão à mediana poderia efetuar melhores estimativas. Então Edgeworth (1888), relaxando a restrição de que a soma dos resíduos seja igual a zero ($\sum \hat{u}_i = 0$) (PORTNOY; KOENKER, 1997, p. 281), propõe o primeiro o protótipo do que viria se tornar, na década de 1940, o algoritmo simplex, capaz de obter, de forma iterativa, o intercepto e o coeficiente angular da reta do método dos mínimos desvios absolutos.

Na década de 40 surgiram os primeiros algoritmos simplex destinados à otimização, algoritmos estes que se ajustam às necessidades do métodos dos mínimos desvios absolutos, que não possui solução analítica, mas iterativa. A primeira aplicação dos algoritmos simplex para a resolução do método dos mínimos desvios absolutos se deve a Charnes, Cooper e Ferguson (1955) (ver PORTNOY; KOENKER, 1997, p. 281; KOENKER, 2018, p. 4).

Barrodale e Roberts (1974) propuseram a forma moderna do algoritmo simplex que, por muitos anos e até hoje é utilizado para a minimização do erro médio absoluto.

Koenker e Gilbert Bassett (1978) generalizaram o problema de minimização do erro médio absoluto, o que equivale à regressão à mediana, ao problema de encontrar os

¹Gauss (1809) ligou o método dos mínimos quadrados à distribuição normal mas a origem do método se deve ao trabalho pioneiro de Legendre (1805). Houve discordâncias entre os dois na disputa pela invenção do método, entre outros achados na época. Ver a este respeito Stigler (1977); Stigler (1981) e Stigler (1986).

diversos quantis de distribuição através da aplicação de uma função de perda assimétrica, correspondente àquele quantil, chegando-se assim à regressão quantílica.

2.2 Referencial teórico

2.2.1 Interpretação Geométrica

2.2.1.1 O gráfico dual

[Edgeworth \(1888\)](#) criou o gráfico dual, ferramenta essencial para o desenvolvimento do seu procedimento de minimização dos resíduos absolutos. O gráfico dual nada mais é do que o gráfico, em duas dimensões, dos parâmetros α e β a serem estimados, com o parâmetro α nas ordenadas e o parâmetro β no eixo das abscissas.

2.2.1.2 Algoritmo inicial

[Edgeworth \(1888\)](#) (*apud Koenker (2000)*, p.356) mostrou que, para a regressão à mediana bivariada, pontos no espaço amostral correspondem a retas no espaço paramétrico $(x_i, y_i) \mapsto \{(\alpha, \beta) : \alpha = y_i - x_i\beta\}$, e retas através de pares de pontos no espaço amostral correspondem a pontos (intersecção das duas retas geradas pelos pontos) no espaço paramétrico. Todos os pares de observações assim gerados produzem $\binom{n}{2}$ pontos no espaço paramétrico. Num espaço de p dimensões, o número de vértices é da ordem $\binom{n}{p} = \mathcal{O}(n^p)$.

[Edgeworth \(1888\)](#) então propôs um algoritmo geométrico, iniciando em algum ponto do espaço paramétrico, mas especificamente em um ponto de intersecção de duas retas e, iterativamente, computando a derivada direcional da função erro absoluto (ℓ_1), até encontrar uma solução ótima, ou seja, até se chegar a um ponto em que as derivadas direcionais não sejam mais descendentes, ou seja, não haja mais nenhuma direção onde o erro absoluto esteja diminuindo (ver [KOENKER, 2000](#), p. 356-358).

O procedimento acima pode ser algebraicamente escrito e generalizado para n dimensões.

2.2.2 O problema de estimação de quantis como um problema de minimização

Enquanto a regressão linear, também chamada de estimador dos mínimos (erros) quadrados, é o método análogo, para n dimensões, da média amostral, o estimador dos mínimos erros absolutos é o método análogo à mediana amostral em n dimensões ([BASSETT, G.; KOENKER, 1978](#), p. 618).

Desta maneira, é útil definir algebraicamente a média μ e a mediana amostral Me , como minimizações de um tipo de erro, quadrático ou absoluto. A regressão linear e a quantílica, então, são facilmente generalizadas pela extensão destas minimizações unidimensionais, para n dimensões.

Pode-se demonstrar que, assim como a média aritmética μ de uma variável aleatória tem a propriedade de minimizar a soma dos desvios quadráticos de cada observação em relação a ela ([MATLOFF, 2017](#), p. 50), a mediana tem a propriedade de minimizar a soma dos desvios médios absolutos de cada observação ([MATLOFF, 2017](#), p. 260). Ou seja:

$$\mu(Y) = \mathbb{E}Y = \arg \min_c \sum_{i=1}^n \frac{1}{n} (y_i - c)^2$$

$$Me = \arg \min_c \sum_{i=1}^n \frac{1}{n} |y_i - c|$$

Sabe-se que a mediana de uma variável equivale ao quantil de 50%. Assim, outros quantis podem ser obtidos com formulação análoga à formulação acima, porém com a aplicação de uma função de perda assimétrica ($\rho_\tau(\cdot)$) em lugar da função módulo (ver figura 4):

$$Q_\tau(Y) = \arg \min_c \sum_{i=1}^n \rho_\tau(y_i - c)$$

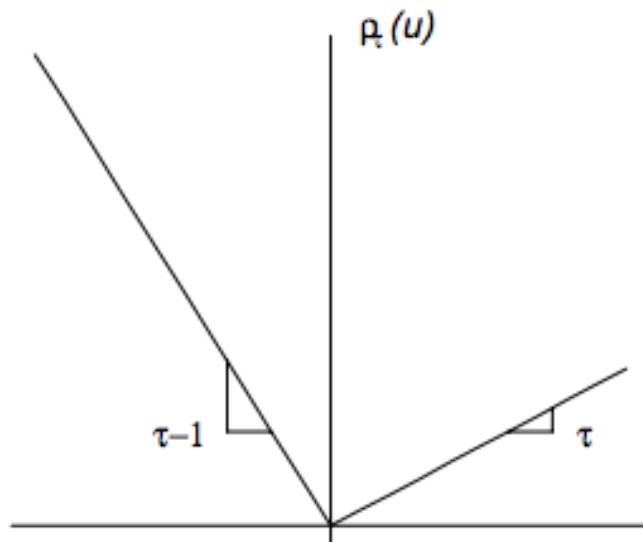


Figura 4: Função de perda ou custo.

Fonte: [Koenker e Hallock \(2001\)](#).

2.2.3 Regressão linear e quantílica

A regressão linear pode ser vista como uma forma de minimização, assim como a média de uma população pode ser visto como o problema de minimização descrito acima.

A diferença é que no caso da regressão linear, ao invés de minimizar em relação a um escalar, desta vez se minimiza o erro em prever uma variável Y em relação a uma função de outra variável X , $f(X)$. Pode-se demonstrar que entre todas as funções $f(X)$, a que minimiza o erro médio quadrático de Y dado X ($\mathbb{E}[(Y - f(X))^2]$) é a função de regressão $\mu(t) = \mathbb{E}(Y|X = t)$ ([MATLOFF, 2017](#), p. 49-50). Logo, estima-se $\hat{\beta}$ assim:

$$\hat{\beta} = \arg \min_{b \in \mathbb{R}^p} \sum (y_i - x_i' b)^2$$

Analogamente, pode-se demonstrar que a mediana condicional é a função que minimiza o erro médio absoluto de Y dado X ($\mathbb{E}(|Y - f(X)|)$) (MATLOFF, 2017, p. 260-261).

Desta maneira, na regressão quantílica estima-se os coeficientes de acordo com o seguinte problema de minimização (KOENKER, 2017, p. 2):

$$\hat{\beta}(\tau) = \arg \min_{b \in \mathcal{R}^p} \sum \rho_\tau(y_i - x'_i b)$$

Esta solução tem a propriedade de que, aproximadamente, τn são positivos e $(1 - \tau)n$ são negativos (KOENKER, 2017, p. 3).

No caso particular da regressão à mediana, essa equação se reduz à:

$$\hat{\beta}(\tau = 0,5) = \arg \min_{b \in \mathcal{R}^p} \sum |y_i - x'_i b|$$

A diferença, portanto, da regressão linear à média e da regressão à mediana, pode ser escrita com a diferença da definição da norma de um espaço vetorial, aplicada ao vetor de erros $\varepsilon = y_i - x'_i \beta$.

No caso da regressão à mediana, essa norma é denominada norma ℓ_1 , enquanto na regressão à média, a norma é denominada norma ℓ_2 , ou norma euclidiana, de maneira que o problema de minimização pode ser genericamente escrito como a seguir:

$$\arg \min_{b \in \mathcal{R}^p} \sum \|y_i - x'_i b\|$$

Na regressão à mediana, a norma de um vetor x é definida da seguinte maneira:

$$\|x\|_1 = \sum_i |x_i| = |x_1| + |x_2| + \dots + |x_i|$$

Na regressão linear, conforme afirmou-se, a norma utilizada é a euclidiana, assim definida:

$$\|x\|_2 = \sqrt{\left(\sum_i x_i^2 \right)} = \sqrt{x_1^2 + x_2^2 + \dots + x_i^2}$$

2.2.3.1 Unicidade da solução

Pode-se demonstrar que a regressão linear, ou seja, a minimização de $\mathbb{E}[(Y - X\beta)^2]$ possui uma única solução e esta solução pode ser encontrada analiticamente, bastando para isso efetuar a derivação parcial deste termo em relação a β e igualando-o a zero, obtendo-se assim uma única solução para β a qual usualmente designa-se $\hat{\beta}$ (ver MATLOFF, 2017, p. 49-50).

O mesmo não se pode dizer da regressão à mediana e, mais genericamente, da regressão quantílica. Nesta abordagem, há múltiplas soluções possíveis, assim como numa amostra de tamanho par existem duas medianas possíveis. Ainda, as soluções do problema

de minimização da regressão quantílica não podem ser encontradas analiticamente, sendo necessária a utilização de processos iterativos para a obtenção do(s) mínimo(s).

Contudo, deve-se ter em mente que, em ambos os processos de minimização, seja para a regressão linear ou para a regressão quantílica, trabalha-se com apenas uma amostra da população estudada. Desta forma, os valores de $\hat{\beta}$ encontrados são apenas estimativas dos valores reais de β , ou seja, os valores da população.

A diferença entre as múltiplas soluções da regressão quantílica é da ordem de $1/n$, enquanto a amplitude da precisão da estimativa é de tamanho $1/\sqrt{n}$. Assim, presume-se que as múltiplas soluções possíveis, para os casos práticos estão dentro da margem de erro para a primeira estimativa encontrada pelo algoritmo ([BASSETT, G.; KOENKER, 1978](#)).

2.2.3.2 Robustez da solução

Uma das principais vantagens da regressão à mediana consiste na sua robustez à presença de *outliers*. Enquanto a solução da regressão linear é altamente sensível à presença de eventuais *outliers* na amostra, a solução de regressão quantílica é muito menos sensível.

A figura 5a² ilustra o caso típico de um *outlier* presente na variável resposta do modelo, onde a solução da regressão quantílica se mostra totalmente estável, pois este ponto não altera substancialmente a estimativa (reta em vermelho) da reta de regressão quantílica “real” (em azul). Já a figura 5b³ ilustra um caso onde o *outlier* é um ponto discrepante também na variável explicativa, situação esta que não é incomum em amostras de imóveis urbanos. Neste último caso, a solução da regressão quantílica mostra-se totalmente sensível à influência deste ponto.

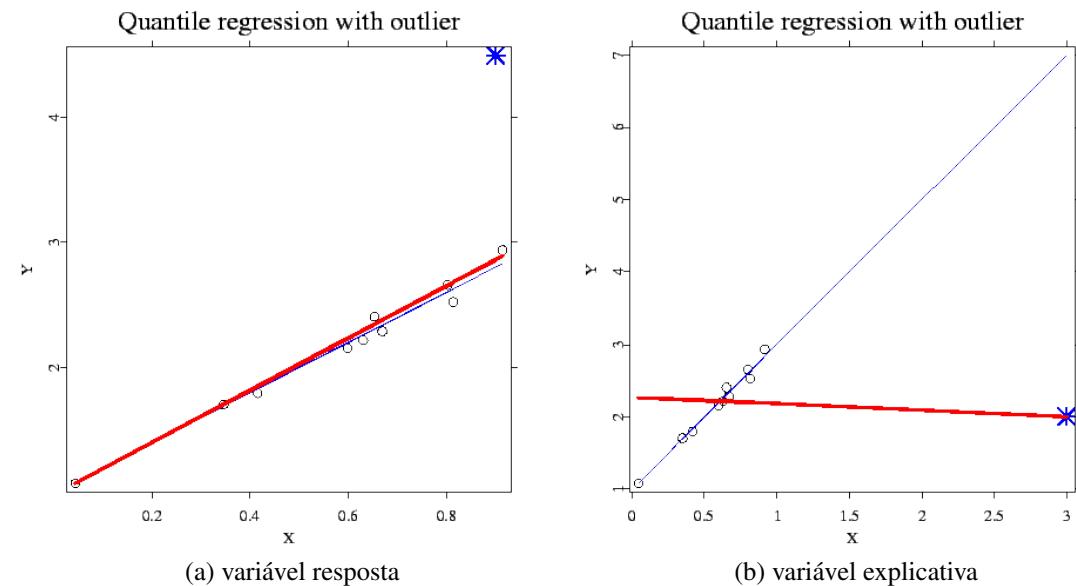


Figura 5: Regressão quantílica: efeito de outliers na amostra.

²Fonte: http://sfb649.wiwi.hu-berlin.de/fedc_homepage/xplore/tutorials/xagtmlnode6.html

³Idem.

2.2.3.3 Transformação e retransformação

Na regressão linear, com a aplicação do método à uma variável produto da transformação da variável dependente original, não é suficiente a simples retransformação dessa variável à escala original para a obtenção da solução, já que $f(\mathbb{E}(X)) \leq \mathbb{E}(f(X))$, devido à desigualdade de Jensen (ver DROUBI; HOCHHEIM; ZONATO, 2019, p. 207).

Já na regressão quantílica, com a aplicação de **transformações monotônicas**, pode-se afirmar que o quantil da variável transformada corresponde ao mesmo quantil da variável original. Assim, seja $f(\cdot)$ uma transformação monotônica qualquer em \mathbb{R} , e $Q(\cdot)$ a função quantil, para uma variável aleatória qualquer Y pode-se escrever:

$$Q_{f(Y)}(\tau) = f(Q_Y(\tau))$$

Koenker e Gilbert Bassett (1978, p. 39-40) elencam ainda quatro propriedades de equivariância para a regressão quantílica, reproduzidas abaixo:

$$\begin{aligned}\hat{\beta}(\tau; \lambda y, X) &= \lambda \hat{\beta}(\tau; y, X), & \lambda \in [0, \infty), \\ \hat{\beta}(1 - \tau; \lambda y, X) &= \lambda \hat{\beta}(\tau; y, X) & \lambda \in (-\infty, 0], \\ \hat{\beta}(\tau; y + X\gamma, X) &= \hat{\beta}(\tau; y, X) + \gamma & \gamma \in \mathbb{R}^k, \\ \hat{\beta}(\tau; y, XA) &= A^{-1} \hat{\beta}(\tau; y, X) & A_{K \times K} \text{ não-singular.}\end{aligned}$$

2.2.3.4 Eficiência computacional

Segundo Koenker (2018, p. 3), para amostras de tamanho menor do que alguns milhares de dados, com algumas dúzias de parâmetros a serem estimados, o que é hoje considerado um problema modesto pela ciência de dados, o algoritmo de Barrodale e Roberts (1974) é considerado extremamente eficiente.

De fato, segundo Koenker (2018, p. 6), algoritmos do tipo simplex como o de **barrolade**, ditos *exterior point algorithms*, são mais eficientes para problemas de número moderado de dados e parâmetros.

Já para a solução de problemas com maior número de dados, procedimentos do tipo *interior point algorithms*, como o de Frisch-Newton, são substancialmente mais rápidos e precisos (KOENKER, 2018, p. 6). De fato, como mencionado anteriormente, com o algoritmo simplex o problema é da ordem $\mathcal{O}(n^p)$. Segundo Koenker (2000, p. 369), com a utilização de um *interior point algorithm*, o problema toma a ordem de $\mathcal{O}(np^3 \log^2 n)$, ainda muito superior à ordem do problema de mínimos quadrados, que é de ordem $\mathcal{O}(np^2)$.

Com algum pré-processamento, é possível reduzir bastante a ordem do problema.

Para problemas esparsos, *i.e.* problemas onde há um grande número de variáveis, porém apenas algumas com valores diferentes de zero, foram desenvolvidos outros métodos baseados em álgebra esparsa (ver KOENKER; NG, 2005).

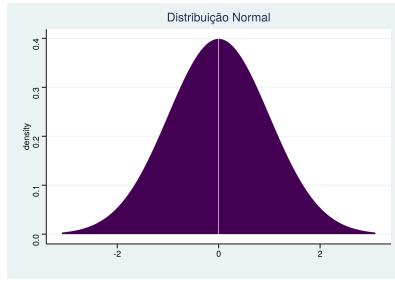
Uma comparação detalhada da eficiência dos diversos métodos computacionais com diferentes tamanhos de amostras e parâmetros pode ser visto em Portnoy e Koenker (1997).

Mais recentemente, com o advento do *Big Data*, mesmo os algoritmos do tipo *interior point algorithms*, ainda que com uso de pré-processamento, podem ser insuficientes. Para isto, é necessário fazer uso do método de gradiente descendente, com o auxílio de computação paralela ([KOENKER; HALLOCK, 2001](#), p. 8).

2.2.3.5 Estimador de máxima verossimilhança

Pode-se demonstrar que, quando a distribuição é normal, o estimador de máxima verossimilhança para o parâmetro μ da distribuição é a média amostral.

Analogamente, se a distribuição dos dados for a distribuição de Laplace, o estimador de máxima verossimilhança para o parâmetro é a mediana.



$$\hat{\mu} = \frac{1}{n} \sum x_i$$

$$\hat{\sigma} = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \hat{\mu})^2}$$

$$f(x|\mu, \sigma) = \frac{1}{\sigma \sqrt{2/\pi}} \exp\left(-\frac{1}{2} \frac{(x-\mu)^2}{\sigma^2}\right)$$

Figura 6: Distribuição Normal.

$$\hat{\mu} = \arg \min_c \sum |x_i - c|$$

$$\hat{\lambda} = \frac{1}{n} \sum_{i=1}^n |x_i - \hat{\mu}|$$

$$f(x|\mu, \lambda) = \frac{1}{2\lambda} \exp\left(-\frac{|x-\mu|}{\lambda}\right)$$

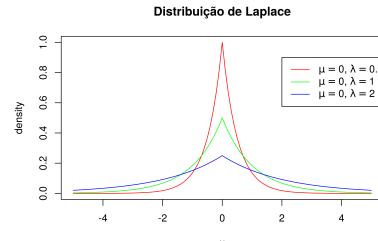


Figura 7: Distribuição de Laplace.

2.2.3.6 Eficiência estatística

Como mostrado no item anterior, a média é o estimador de máxima verossimilhança quando os dados tem distribuição normal, enquanto a mediana é o estimador de máxima verossimilhança quando os dados assumem a distribuição de Laplace.

Na prática, no entanto, os dados não são perfeitamente normais ou perfeitamente laplacionas, porém ajustam-se melhor ou pior a uma ou outra, ou ainda à soma de diversas distribuições de probabilidade.

Pode-se demonstrar que, se por um lado a média amostral tem eficiência estatística assintótica igual a 1 quando a distribuição dos dados é normal ou gaussiana, porém ela tem menos da metade da eficiência da mediana quando a distribuição for a distribuição de Laplace e tem eficiência zero para a distribuição de Cauchy ([KOENKER; BASSETT, G., 1978](#), p.36).

Isto implica que, se a distribuição dos resíduos é normal, são necessários $\pi/2$ (1,57) vezes mais dados para que a estimativa de μ através da mediana seja tão eficiente quanto a estimativa através da média. Logo, neste caso, os intervalos de confiança obtidos para a regressão quantílica são 25% mais largos do que os intervalos de confiança para a regressão linear (KOENKER, 2000, p. 354; DASGUPTA, 2008, p. 92). No entanto, se a distribuição dos dados for a distribuição de Laplace, são necessários duas vezes mais dados para que a média estime μ com a mesma precisão da mediana.

Desta maneira, quando há incerteza quanto à real distribuição dos dados, a média, no caso escalar, ou a média condicional (regressão linear), considerada o estimador ótimo para o caso gaussiano, pode ser preterível à mediana, ou outros estimadores, ditos “ineficientes” (KOENKER; BASSETT, G., 1978, p.36).

A hipótese da normalidade dos dados é virtualmente preestabelecida na comunidade científica em geral. No entanto, a verificação da normalidade nem sempre é feita de uma maneira rigorosa. Um caso interessante ilustra este fato: ao analisar um experimento desenhado por Pierce (1932) em 1872 para investigar a validade da lei dos erros de Gauss, assim como a aplicabilidade do método dos mínimos quadrados, em que este concluiu pela justificativa da adoção da hipótese da normalidade, Fréchet (1924) e depois Wilson e Hilferty (1929) concluíram que os dados se adaptavam melhor à primeira lei dos erros de Laplace, $\varphi(\varepsilon) = Ce^{-|Re|}$, do que à segunda, $\varphi(\varepsilon) = Ce^{-R^2\varepsilon^2}$, i.e. à distribuição de Gauss (KOENKER, 2009, p. 62).

Segundo Gilbert Bassett e Koenker (1978, p. 618), para qualquer distribuição de erros em que a mediana é superior à média como um estimador de tendência central (isto se aplica a uma enorme classe de distribuições com caudas longas e/ou densidade máxima na mediana), o estimador de mínimos desvios absolutos é preferível ao estimador de mínimos quadrados, apresentando menores elipsóides de confiança assintoticamente.

2.3 Inferência e Qualidade do ajuste da regressão quantílica

Na regressão linear, normalmente a verificação da qualidade do ajuste do modelo é feita através do coeficiente de determinação R^2 , ou do coeficiente de determinação ajustado. Koenker e Machado (1999) mostra como é calculado a qualidade do ajuste dos dados a um modelo de regressão quantílica.

Quando se considera o modelo de regressão linear em que os erros apresentam distribuição normal, o coeficiente de determinação R^2 é uma medida bastante utilizada para análise da qualidade do ajuste do modelo. Para os modelos de regressão quantílica, Koenker e Machado (1999) propuseram uma medida semelhante ao coeficiente de determinação R^2 , que permite ao pesquisador avaliar a qualidade do ajuste do modelo. Para essa análise, considere-se um modelo linear para o quantil condicional de $(y_i|x_i)$, com p variáveis explicativas. Teremos:

$$Q(y_i|x_i) = x_i' \beta(\tau) = x_{i1}' \beta_1(\tau) + x_{i2}' \beta_2(\tau)$$

Em que x_i , i -ésima linha da matriz de planejamento X , é partitionada em duas partes, aqui chamadas de x_{i1} e x_{i2} de dimensões $p \times q$ e q , respectivamente. Considere-se que $\hat{\beta}(\tau)$ é

o estimador responsável por minimizar a soma dos desvios absolutos ponderados para o modelo completo. Então teremos:

$$\hat{V}(\tau) = \sum_{i=1}^n \rho_\tau[y_i - x'_i \beta(\tau)]$$

E $\tilde{\beta}(\tau)$, sob restrição q-dimensional $H_0 : \beta_2(\tau) = 0$, o estimador responsável por minimizar a soma dos desvios absolutos ponderados para o modelo reduzido:

$$\tilde{V}(\tau) = \sum_{i=1}^n \rho_\tau[y_i - x'_i \tilde{\beta}(\tau)]$$

Dessa forma, o coeficiente de determinação para regressão quantílica do modelo (1.5), pode ser definido da seguinte maneira:

$$R^1 = 1 - \frac{\hat{V}(\tau)}{\tilde{V}(\tau)}$$

Considerando que $\tilde{\beta}(\tau)$ é obtido quando restringimos $\hat{\beta}(\tau)$, então $\tilde{V}(\tau) \geq \hat{V}(\tau)$ e assim tem-se que $R^1(\tau)$ terá valores dentro do intervalo $R^1(\tau) \in [0, 1]$, sendo desta forma uma medida possível para mensurar a qualidade do ajuste para um modelo indexado a um determinado quantil da regressão quantílica.

Se considerarmos no vetor de parâmetros $\beta_2(\tau)$ os coeficientes de regressão associados a todas as variáveis explicativas disponíveis, de forma que o modelo reduzido tenha apenas o intercepto, então $R^1(\tau)$ calculado se assemelha bastante ao coeficiente de explicação R^2 comumente utilizado na análise de regressão clássica ([SANTOS, 2012](#)).

Na medida em que o mede o relativo sucesso de dois modelos para a média condicional em função de termos da variância residual, [Koenker e Machado \(1999\)](#) (apud [Santos \(2012\)](#)), nos mostram que o $R^1(\tau)$ mede o relativo sucesso de correspondentes modelos de regressão quantílica em um específico quantil em função de umapropriada soma de resíduos absolutos ponderados. Sendo assim, o $R^1(\tau)$ trata-se de uma medida local de qualidade de ajuste do modelo de regressão quantílica para um determinado quantil.

É de se esperar que o valor deste critério esteja está entre 0 e 1. Quanto maior o coeficiente de determinação $R^1(\tau)$, melhor a qualidade do modelo ajustado.

A teoria relativa a regressão quantílica já evoluiu bastante desde sua proposta original, com extensões em classes de modelos não lineares propostos por [Koenker e Park \(1996\)](#). O comportamento assintótico dos estimadores dos modelos de regressão quantílica não linear se assemelha à teoria dos mínimos quadrados não lineares, de forma que a inferência para regressão quantílica não linear pode ser adaptada diretamente dos métodos destes ([KOENKER, 2005](#)).

2.3.1 Intervalos de Confiança

Segundo [Koenker e Hallock \(2001\)](#), é um princípio básico da Econometria que qualquer estimador sério deve ter uma forma de verificação da sua precisão. Existem

diversas formas de se construir intervalos de confiança para a regressão quantílica. A descrição destes métodos está fora do escopo deste artigo. Segundo [Koenker e Hallock \(2001, p. 153\)](#), as discrepâncias entre os intervalos obtidos para a regressão quantílica e os outros métodos disponíveis na econometria são pequenos e a inferência para a regressão quantílica é talvez mais robusta do que as outras formas de inferência comumente encontradas.

Uma comparação de diversos métodos de cômputo dos intervalos de confiança para diferentes distribuições de erros pode ser encontrada em [Koenker \(1994\)](#).

2.4 Aplicações da regressão quantílica

Aplicações da regressão quantílica tem aplicação em especial na Econometria, mas também nos mais diversos ramos da ciência, na solução ou descrição de diversos problemas. Para isto, foram desenvolvidos diferentes métodos de análise de quantis, análogos aos variados métodos de regressão existentes para a média.

Para dados em séries temporais, foram desenvolvidos métodos para modelos autoregressivos (AR), autoregressivos heteroscedásticos (ARCH) e ARCH generalizados (GARCH) (ver [KOENKER; XIAO, 2006](#); [KOENKER; ZHAO, 1996](#); [XIAO; KOENKER, 2009](#)). Para dados longitudinais, ou em painel, foram desenvolvidos modelos mistos (ver [KOENKER, 2004](#)).

Aplicações interessantes nas finanças e Econometria incluem a otimização de portfólios de investimentos ([BASSETT G. W., J.; KOENKER; KORDAS, 2004](#)), e análise de dados de desemprego ([BILIAS; KOENKER, 2001](#)).

Um resumo dos diversos métodos e aplicações da regressão quantílica podem ser encontrados em [Koenker \(2017\)](#).

3 Estudos de Caso

Para os estudos de caso foram utilizados os dados disponíveis em [Hochheim \(2015\)](#).

3.1 Duas dimensões

Assim como na regressão linear, é mais fácil a compreensão da regressão quantílica através de exemplos em duas dimensões, e depois generalizar para n dimensões.

Seja primeiramente o caso de dados heteroscedásticos. A figura 8 ilustra a aplicação da regressão quantílica e da regressão linear para este caso. Na figura 8, a reta vermelha é a reta de regressão linear entre as variáveis. A área sombreada em cinza é o intervalo de confiança para a regressão linear @80%. As retas azuis são as retas de regressão quantílica para os decís $D_1 (0,1)$ a $D_9 (0,9)$.

A regressão quantílica neste caso pode ser usada para demonstrar a não validade dos intervalos de confiança (IC) e predição (IP) para a regressão linear para este tipo de dados: como a variância da população não é constante, mas aumenta com o aumento da

área, as retas da regressão quantílica se abrem. Como os intervalos de confiança e predição na inferência clássica são calculados considerando-se que a variância da população é constante, este efeito não se observa no formato do IC.

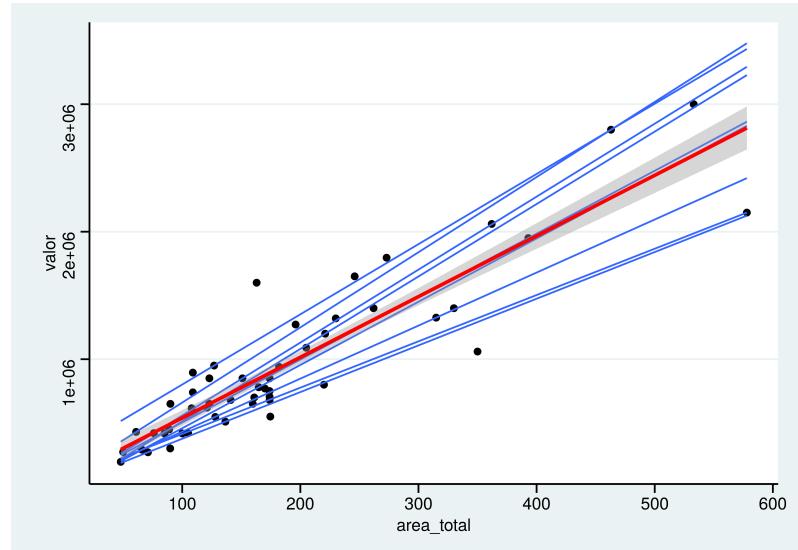


Figura 8: Regressão Linear e Quantílica para dados heteroscedásticos.

Assim como na regressão linear, uma conveniente transformação das variáveis pode ser aplicada para a obtenção da homoscedasticidade. Isto pode ser visto na figura 9, onde as retas para os diferentes quantis obtidas pela regressão quantílica agora são praticamente paralelas entre si, indicando que a heteroscedasticidade foi removida.

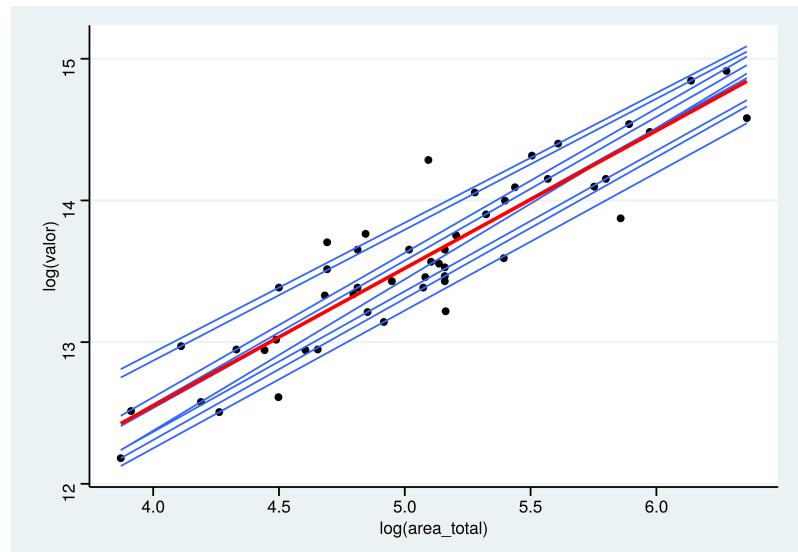


Figura 9: Regressão Linear e Quantílica com dados transformados.

Os coeficientes das retas de regressão quantílica podem ser plotados como na figura 10. Nesta figura, a reta cheia vermelha representa o coeficiente do modelo de regressão linear, enquanto a reta preta pontilhada representa os vários coeficientes da regressão quantílica. As retas vermelhas tracejadas representam o intervalo de confiança

de estimação do coeficiente de regressão linear. A área sombreada em cinza representa os intervalos de confiança para os coeficientes da regressão quantílica. Deve-se notar que, entre os quantis aproximados de 0,3 e 0,55, os coeficientes da regressão quantílica não são significantemente diferentes, estatisticamente, do coeficiente da regressão linear.

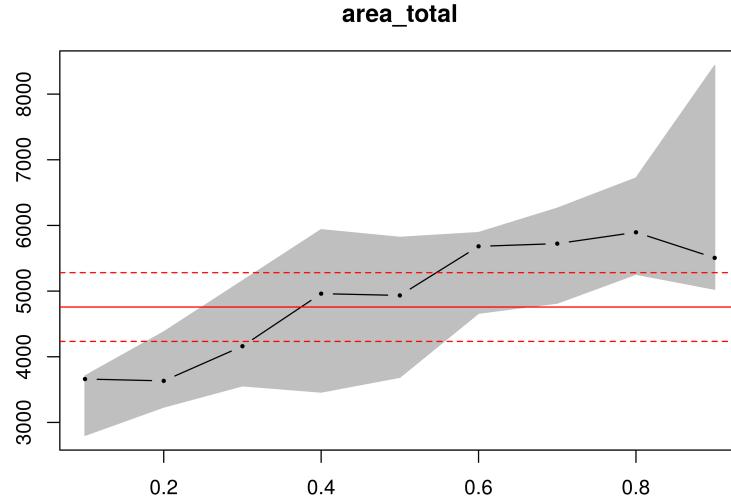


Figura 10: Variação dos coeficientes de regressão quantílica (variáveis originais).

Outra forma de se analisar este gráfico é: a área total tem maior influencia na formação de valor de um imóvel nos quantis superiores (acima de 0,55) e menor influência nos quantis inferiores (abaixo de 0,3). Os coeficientes para os quantis 0,3; 0,5 e 0,8 podem ser vistos na tabela 1. Para um imóvel no quantil central ($\tau = 0,5$), uma unidade marginal de área adiciona aproximadamente 18,5% mais valor do que em um imóvel nos quantis inferiores e para um imóvel nos quantis superiores uma unidade marginal de área adiciona aproximadamente 19,5% mais valor do que para um imóvel no quantil central ($\tau = 0,5$).

Está claro, aqui, que o efeito maior da variável área total nos quantis superiores é devido ao efeito de um covariante, por exemplo, o padrão de acabamento, variável omitida do modelo, que deve estar correlacionada com a variável área total (imóveis de alto padrão usualmente contam com área total maior e imóveis de menor área tem menor padrão de acabamento). Mas isto serve para ilustrar bem o funcionamento da regressão quantílica e uma de suas vantagens em relação à regressão linear: na regressão linear, a omissão da variável padrão de acabamento, em conjunto com um único coeficiente médio para a variável área resultaria em um modelo viesado. Já na regressão quantílica, mesmo a variável padrão de acabamento estando ausente, o seu efeito se faz sentir pela mudança na magnitude do coeficiente da variável presente, ou seja, na variável área total.

Já no caso dos dados transformados, pode-se notar na figura 11 que para todos os quantis, os coeficientes da regressão quantílica não podem ser considerados estatisticamente diferentes do coeficiente da regressão linear. Também se pode notar nesta figura como o estimador de regressão linear, para uma variável normalmente distribuída e na ausência de heteroscedasticidade, é mais eficiente do que o estimador da regressão quantílica, como a teoria já prevê (ver Matloff (2017), 238).

Tabela 1: Comparação entre diferentes quantis da regressão quantílica.

	<i>Dependent variable:</i>		
	valor		
	$\tau = 0,3$	$\tau = 0,5$	$\tau = 0,8$
Constant	15.397,59 (63.253,00)	10.855,26 (69.056,43)	70.373,13 (102.793,74)
	$t = 0,24$	$t = 0,16$	$t = 0,68$
	$p = 0,81$	$p = 0,88$	$p = 0,50$
area_total	4.160,64 (571,63)	4.934,21 (565,46)	5.895,52 (407,86)
	$t = 7,28$	$t = 8,73$	$t = 14,45$
	$p = 0,00^{***}$	$p = 0,00^{***}$	$p = 0,00^{***}$
Observations	50	50	50

Note: * $p < 0,3$; ** $p < 0,2$; *** $p < 0,1$

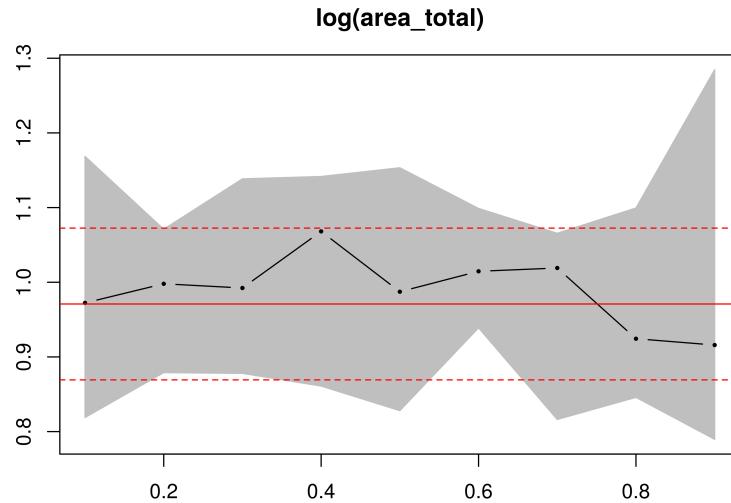


Figura 11: Variação dos coeficientes de regressão quantílica (variáveis transformadas).

3.2 Análise Multivariada

Para os dados obtidos de Hochheim (2015, p. 22-23) foram ajustados dois modelos, um de regressão linear, com os dados saneados, e outro de regressão quantílica, utilizando-se a totalidade dos dados, para os decis D_1 (0,1) a D_9 (0,9).

Na figura 12 podem ser vistos os valores dos coeficientes de cada variável para os diferentes quantis. Pode-se perceber, mais uma vez, que o valor dos coeficientes da regressão quantílica não diferem significantemente dos coeficientes da regressão linear (exceção para alguns quantis superiores nas variáveis area_total e padrao).

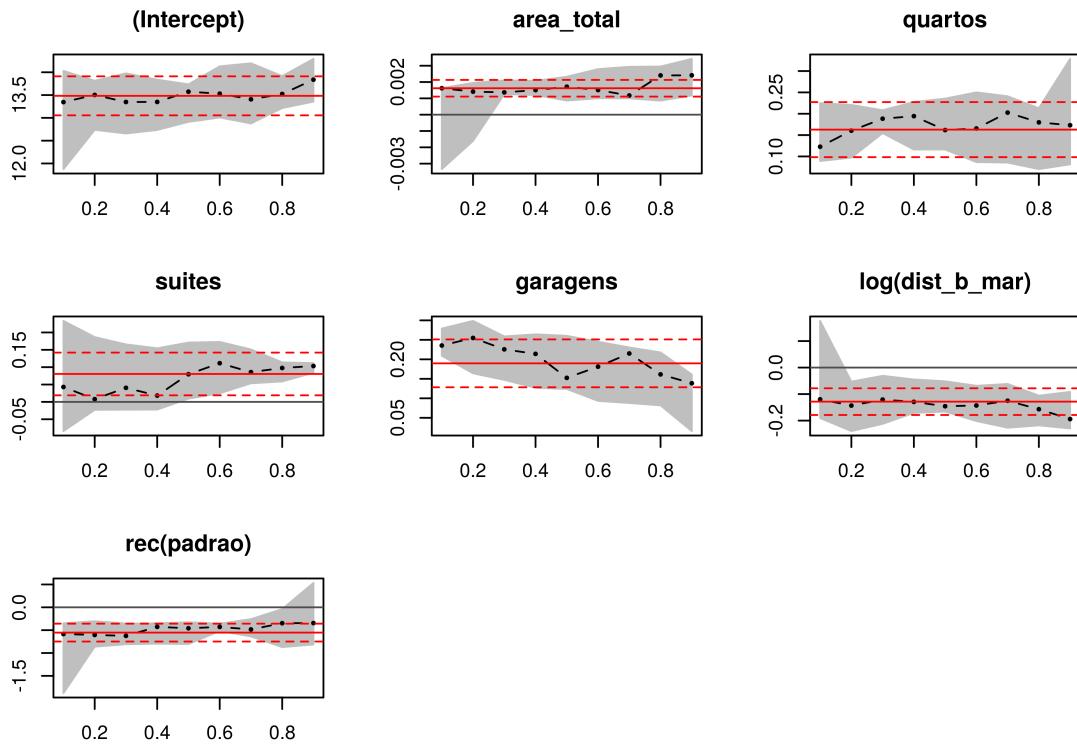


Figura 12: Coeficientes de regressão linear e quantílica. Análise multivariada.

Na tabela 2 podem ser vistos os coeficientes e estatísticas básicas dos modelos de regressão linear e de regressão à mediana (quantil 0,5).

3.2.1 Estimativas

É interessante comparar as estimativas obtidas com os modelos de regressão linear, com dados saneados, e o modelo de regressão à mediana, com a totalidade dos dados. Por um lado, o modelo de regressão linear tende a ser mais preciso para a estimativa da média, como prevê a teoria. Por outro lado, com mais dados, o modelo de regressão à mediana pode tornar-se mais eficiente.

Deve-se levar em conta que as estimativas com o modelo de regressão linear aqui apresentadas são para a mediana da distribuição lognormal.

Pelo modelo de regressão linear, o valor da estimativa central encontrado foi de R\$961.660,64, com intervalo de confiança entre R\$ 924.768,13 e R\$ 1.000.024,94. A amplitude do intervalo de confiança foi de 8%.

Já pelo modelo de regressão quantílica, o valor da estimativa central encontrado foi de R\$946.467,87, com intervalo de confiança entre R\$ 886.472,34 e R\$ 1.010.523,85. A amplitude do intervalo de confiança foi de 13%.

O modelo de regressão linear mostrou-se, portanto, mais eficiente do que o modelo de regressão à mediana, apesar no menor número de dados. Os limites inferior e superior do intervalo de predição @80% para o modelo de regressão linear são, respectivamente: R\$ 802.017,63 e R\$ 1.153.080,88.

Tabela 2: Comparação entre os modelos de regressão linear e regressão à mediana.

	<i>Dependent variable:</i>	
	<i>OLS</i>	<i>quantile regression</i>
	(1)	(2)
Constant	13,564 (13,268, 13,859) t = 58,847 p = 0,000***	13,574 (13,100, 14,047) t = 36,732 p = 0,000***
area_total	0,001 (0,001, 0,002) t = 5,113 p = 0,00001***	0,002 (0,001, 0,003) t = 2,300 p = 0,027***
quartos	0,164 (0,118, 0,209) t = 4,626 p = 0,00004***	0,162 (0,107, 0,217) t = 3,788 p = 0,0005***
suites	0,061 (0,018, 0,104) t = 1,810 p = 0,078***	0,080 (0,020, 0,139) t = 1,712 p = 0,095***
garagens	0,209 (0,166, 0,252) t = 6,247 p = 0,00000***	0,152 (0,075, 0,230) t = 2,520 p = 0,016***
log(dist_b_mar)	-0,141 (-0,176, -0,106) t = -5,174 p = 0,00001***	-0,146 (-0,210, -0,081) t = -2,904 p = 0,006***
rec(padrao)	-0,563 (-0,697, -0,428) t = -5,360 p = 0,00001***	-0,459 (-0,650, -0,267) t = -3,070 p = 0,004***
Observations	48	50
R ²	0,956	
Adjusted R ²	0,950	
Residual Std. Error	0,136 (df = 41)	
F Statistic	148,921*** (df = 6; 41)	

Note:

*p<0,3; **p<0,2; ***p<0,1

20 / 25

Para o modelo de regressão quantílica, o intervalo de predição não faz qualquer sentido. No entanto, é possível estimar os valores diretamente para os decis D_1 e D_9 da população, para efeito de comparação com os limites do intervalo de predição para a regressão linear. Nesta caso, os valores encontrados foram, respectivamente: R\$ 810.629,32 e R\$ 1.186.954,14.

Podem ainda ser calculados os intervalos de confiança @80% para as estimativas dos decis D_1 e D_9 .

Os limites inferior e superior do IC para o D_1 são, respectivamente: R\$ 781.253,06 e R\$ 841.110,17.

Os limites inferior e superior do IC para o D_9 são, respectivamente: R\$ 1.116.547,53 e R\$ 1.261.800,41.

A tabela abaixo resume os resultados obtidos acima: pode-se notar que há uma melhor estimação com o método de regressão linear (IC menores).

Tabela 3: Comparação dos modelos de regressão linear e regressão quantílica.

Método	10% / IP	IC inferior	Estimativa Central	IC Superior	90% / IP
Regressão linear	802.018	924.768	961.661	1.000.025	1.153.081
Regressão quantílica	810.629	886.472	946.468	1.010.524	1.186.954

4 Conclusão

A regressão quantílica é um método que deve ser considerado na Engenharia de Avaliações, por possibilitar uma melhor descrição de todo o mercado imobiliário, para além apenas do efeito médio obtido com a regressão linear.

No entanto, deve ser levado em conta a distribuição dos erros do modelo: quando os erros obedecem à lei de Gauss (ou segunda lei de Laplace), *i.e.* se os erros tem distribuição normal, como no estudo de caso dos dados multivariados, onde se obteve normalidade através da simples transformação da variável resposta pela função logarítmica, o método da regressão linear é mais eficiente, até para a estimação da mediana.

Porém, deve-se levar em conta que nem sempre os dados amostrais vão se apresentar tão bem comportados. A presença de heteroscedasticidade pode reduzir significantemente a eficiência da estimação pelo método dos mínimos quadrados. A ausência de normalidade também pode vir a ocorrer. Nem sempre se pode obter a normalidade através de uma simples transformação de variáveis. Além do mais, transformações de variáveis distorcem o modelo, por isso são indesejáveis.

O Engenheiro de Avaliações deve estar preparado para o tratamento dos mais diversos tipos de dados. O método da regressão quantílica pode ser um aliado do Engenheiro de Avaliações quando a estrutura dos dados não for a ideal, ou seja, a estrutura de erros normais e homoscedásticos, onde a regressão linear impera.

Referências

- ABNT. **NBR 14653-2**: Avaliação de Bens – Parte 2: Imóveis Urbanos. Rio de Janeiro, fev. 2011. p. 4.
- BARRODALE, I.; ROBERTS, F. D. K. Solution of an Overdetermined System of Equations in the L1 Norm [F4]. **Commun. ACM**, ACM, New York, NY, USA, v. 17, n. 6, p. 319–320, jun. 1974. ISSN 0001-0782. DOI: [10.1145/355616.361024](https://doi.acm.org/10.1145/355616.361024). Disponível em: <<http://doi.acm.org/10.1145/355616.361024>>.
- BASSETT, Gilbert; KOENKER, Roger W. Asymptotic Theory of Least Absolute Error Regression. **Journal of the American Statistical Association**, [American Statistical Association, Taylor & Francis, Ltd.], v. 73, n. 363, p. 618–622, 1978. ISSN 01621459. Disponível em: <<http://www.jstor.org/stable/2286611>>.
- BASSETT GILBERT W., Jr.; KOENKER, Roger W; KORDAS, Gregory. Pessimistic Portfolio Allocation and Choquet Expected Utility. **Journal of Financial Econometrics**, v. 2, n. 4, p. 477–492, set. 2004. ISSN 1479-8409. DOI: [10.1093/jjfinec/nbh023](https://doi.org/10.1093/jjfinec/nbh023). eprint: <http://oup.prod.sis.lan/jfec/article-pdf/2/4/477/2673928/nbh023.pdf>. Disponível em: <<https://doi.org/10.1093/jjfinec/nbh023>>.
- BILIAS, Yannis; KOENKER, Roger W. Quantile regression for duration data: A reappraisal of the Pennsylvania Reemployment Bonus Experiments. **Empirical Economics**, v. 26, n. 1, p. 199–220, 2001. Disponível em: <<https://EconPapers.repec.org/RePEc:spr:empeco:v:26:y:2001:i:1:p:199-220>>.
- BOSCOVICH, Roger Joseph. *Voyage astronomique et géographique, dans l'état de l'église*. Paris: N. M. Tilliard., 1770.
- CADE, Brian S.; NOON, Barry R. A Gentle Introduction to Quantile Regression for Ecologists. **Frontiers in Ecology and the Environment**, Ecological Society of America, v. 1, n. 8, p. 412–420, 2003. ISSN 15409295. Disponível em: <<http://www.jstor.org/stable/3868138>>.
- CHARNES, A.; COOPER, W. W.; FERGUSON, R. O. Optimal Estimation of Executive Compensation by Linear Programming. **Management Science**, v. 1, n. 2, p. 138–151, 1955. DOI: [10.1287/mnsc.1.2.138](https://doi.org/10.1287/mnsc.1.2.138). eprint: <https://doi.org/10.1287/mnsc.1.2.138>. Disponível em: <<https://doi.org/10.1287/mnsc.1.2.138>>.
- DASGUPTA, Anirban. *Asymtotic Theory of Statistics and Probability*. [S.l.]: Springer, 2008. DOI: [10.1007/978-0-387-75971-5](https://doi.org/10.1007/978-0-387-75971-5).
- DROUBI, Luiz Fernando Palin; HOCHHEIM, Norberto; ZONATO, Willian. Estudos Interdisciplinares nas Ciências Exatas e da Terra e Engenharias. In: Florianópolis: Atena Editora, set. 2019. Avaliação pela moda, média ou mediana? DOI: [10.22533/at.ed.218191109](https://doi.org/10.22533/at.ed.218191109). Disponível em: <<http://droubi.me/sobre2018>>.
- EDGEWORTH, Francis Ysidro. On observations relating to several quantities. **Hermathena**, Trinity College Dublin, v. 6, n. 13, p. 279–285, 1887. ISSN 00180750. Disponível em: <<http://www.jstor.org/stable/23036355>>.
- _____. XXII. On a new method of reducing observations relating to several quantities. **The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science**, Taylor & Francis, v. 25, n. 154, p. 184–191, 1888. DOI: [10.1080/14786448808628170](https://doi.org/10.1080/14786448808628170). eprint: <https://doi.org/10.1080/14786448808628170>. Disponível em: <<https://doi.org/10.1080/14786448808628170>>.

- FRÉCHET, M. Sur la loi des erreurs d'observation. **Matematichiskii Sbornik**, v. 32, p. 5–8, 1924.
- GALTON, F. **Natural Inheritance**. London: MacMillan, 1889.
- GAUSS, Carl Friedrich. **Theoria motus corporum coelestium in sectionibus conicis solem ambientium**. [S.l.: s.n.], 1809.
- HOCHHEIM, Norberto. **Engenharia de Avaliações - Módulo Básico**. Florianópolis: IBAPE - SC, 2015.
- JAMES, Kenneth E. Regression toward the Mean in Uncontrolled Clinical Studies. **Biometrics**, [Wiley, International Biometric Society], v. 29, n. 1, p. 121–130, 1973. ISSN 0006341X, 15410420. Disponível em: <<http://www.jstor.org/stable/2529681>>.
- KOENKER, Roger W. Confidence Intervals for Regression Quantiles. In: _____. **Asymptotic Statistics**. Heidelberg: Physica-Verlag HD, 1994. p. 349–359. ISBN 978-3-642-57984-4.
- _____. Galton, Edgeworth, Frisch, and prospects for quantile regression in econometrics. **Journal of Econometrics**, v. 95, n. 2, p. 347–374, 2000. ISSN 0304-4076. DOI: [https://doi.org/10.1016/S0304-4076\(99\)00043-3](https://doi.org/10.1016/S0304-4076(99)00043-3). Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0304407699000433>>.
- _____. Handbook of Quantile Regression. In: New York: Chapman e Hall/CRC, 2018. Computational Methods for Quantile Regression, p. 13. DOI: [10.1201/9781315120256](https://doi.org/10.1201/9781315120256). Disponível em: <<http://www.econ.uiuc.edu/~roger/research/conopt/computation.pdf>>.
- _____. **Quantile Regression**. [S.l.]: Cambridge University Press, 2005. (Econometric Society Monographs). DOI: [10.1017/CBO9780511754098](https://doi.org/10.1017/CBO9780511754098).
- _____. Quantile regression for longitudinal data. **Journal of Multivariate Analysis**, v. 91, n. 1, p. 74–89, 2004. Special Issue on Semiparametric and Nonparametric Mixed Models. ISSN 0047-259X. DOI: <https://doi.org/10.1016/j.jmva.2004.05.006>. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0047259X04001113>>.
- _____. Quantile Regression: 40 Years On. **Annual Review of Economics**, v. 9, n. 1, p. 155–176, 2017. DOI: [10.1146/annurev-economics-063016-103651](https://doi.org/10.1146/annurev-economics-063016-103651). eprint: <https://doi.org/10.1146/annurev-economics-063016-103651>. Disponível em: <<https://doi.org/10.1146/annurev-economics-063016-103651>>.
- _____. The Median Is the Message: Wilson and Hilmert's Experiments on the Law of Errors. **The American Statistician**, Taylor & Francis, v. 63, n. 1, p. 20–25, 2009. DOI: [10.1198/tast.2009.0004](https://doi.org/10.1198/tast.2009.0004). eprint: <https://doi.org/10.1198/tast.2009.0004>. Disponível em: <<https://doi.org/10.1198/tast.2009.0004>>.
- KOENKER, Roger W; BASSETT, Gilbert. Regression Quantiles. **Econometrica**, v. 46, n. 1, p. 33–50, 1978. Disponível em: <<https://EconPapers.repec.org/RePEc:ecm:emetrp:v:46:y:1978:i:1:p:33-50>>.
- KOENKER, Roger W; HALLOCK, Kevin F. Quantile Regression. **Journal of Economic Perspectives**, v. 15, n. 4, p. 143–156, dez. 2001. DOI: [10.1257/jep.15.4.143](https://doi.org/10.1257/jep.15.4.143). Disponível em: <<http://www.aeaweb.org/articles?id=10.1257/jep.15.4.143>>.

- KOENKER, Roger W; MACHADO, José A. F. Goodness of Fit and Related Inference Processes for Quantile Regression. **Journal of the American Statistical Association**, Taylor & Francis, v. 94, n. 448, p. 1296–1310, 1999. DOI: [10.1080/01621459.1999.10473882](https://doi.org/10.1080/01621459.1999.10473882). eprint: <https://amstat.tandfonline.com/doi/pdf/10.1080/01621459.1999.10473882>. Disponível em: <<https://amstat.tandfonline.com/doi/abs/10.1080/01621459.1999.10473882>>.
- KOENKER, Roger W; NG, Pin T. A frisch-newton algorithm for sparse quantile regression. English (US). **Acta Mathematicae Applicatae Sinica**, Springer Verlag, v. 21, n. 2, p. 225–236, mai. 2005. ISSN 0168-9673. DOI: [10.1007/s10255-005-0231-1](https://doi.org/10.1007/s10255-005-0231-1).
- KOENKER, Roger W; PARK, Beum J. An interior point algorithm for nonlinear quantile regression. **Journal of Econometrics**, v. 71, n. 1, p. 265–283, 1996. ISSN 0304-4076. DOI: [https://doi.org/10.1016/0304-4076\(96\)84507-6](https://doi.org/10.1016/0304-4076(96)84507-6). Disponível em: <<http://www.sciencedirect.com/science/article/pii/0304407696845076>>.
- KOENKER, Roger W; XIAO, Zhijie. Quantile Autoregression. **Journal of the American Statistical Association**, Taylor & Francis, v. 101, n. 475, p. 980–990, 2006. DOI: [10.1198/016214506000000672](https://doi.org/10.1198/016214506000000672). eprint: <https://doi.org/10.1198/016214506000000672>. Disponível em: <<https://doi.org/10.1198/016214506000000672>>.
- KOENKER, Roger W; ZHAO, Quanshui. Conditional Quantile Estimation and Inference for Arch Models. **Econometric Theory**, Cambridge University Press, v. 12, n. 5, p. 793–813, 1996. ISSN 02664666, 14694360. Disponível em: <<http://www.jstor.org/stable/3532912>>.
- LAPLACE, Pierre Simon. Sur quelques points du système du monde. **Mémoires de l'Academie Royale des Sciences de Paris**, p. 1–87, 1793. Reprinted in Laplace, 1878-1912, vol. 11, pp. 477-558.
- LEGENDRE, Adrien Marie. **Nouvelles méthodes pour la détermination des orbites des comètes**. [S.l.]: F. Didot, 1805.
- MATLOFF, Norman. **From Linear Models to Machine Learning**: Regression and Classification, with R examples. [S.l.]: Chapman & Hall, 2017. Disponível em: <<http://heather.cs.ucdavis.edu/draftregclass.pdf>>.
- MOSTELLER, Frederick; TUKEY, John W. **Data Analysis and Regression: a Second Course in Statistics**. [S.l.: s.n.], 1977. p. xvii + 588.
- NEWTON, Isaac. **Philosophiae Naturalis Principia Mathematica**. [S.l.: s.n.], 1687.
- PIERCE, Charles Sanders. The Collected Papers of Charles Sanders Peirce. In: edição: Charles Hartshorne e Paul Weiss. Cambridge: Harvard University Press, 1932. v. 3 On the Theory of Errors of Observation.
- PORTNOY, Stephen; KOENKER, Roger W. The Gaussian Hare and the Laplacian Tortoise: Computability of Squared- Error versus Absolute-Error Estimators. **Statistical Science**, Institute of Mathematical Statistics, v. 12, n. 4, p. 279–296, 1997. ISSN 08834237. Disponível em: <<http://www.jstor.org/stable/2246216>>.
- SANTOS, Bruno Ramos dos. **Modelos de Regressão Quantílica**. 2012. Diss. (Mestrado) – Instituto de Matemática e Estatística, Universidade de São Paulo, São Paulo. DOI: [10.11606/D.45.2012.tde-12042012-154333](https://doi.org/10.11606/D.45.2012.tde-12042012-154333).

STIGLER, Stephen M. An attack on Gauss, published by Legendre in 1820. **Historia Mathematica**, v. 4, n. 1, p. 31–35, 1977. ISSN 0315-0860. DOI: [https://doi.org/10.1016/0315-0860\(77\)90032-5](https://doi.org/10.1016/0315-0860(77)90032-5). Disponível em: <<http://www.sciencedirect.com/science/article/pii/0315086077900325>>.

_____. Gauss and the Invention of Least Squares. **Ann. Statist.**, The Institute of Mathematical Statistics, v. 9, n. 3, p. 465–474, mai. 1981. DOI: [10.1214/aos/1176345451](https://doi.org/10.1214/aos/1176345451). Disponível em: <<https://doi.org/10.1214/aos/1176345451>>.

_____. **The History of Statistics: The Measurement of Uncertainty before 1900**. Cambridge, Mass., & London, England: The Belknap Press of Harvard University Press, 1986. p. 410. ISBN 0674403401.

WILSON, Edwin B.; HILFERTY, Margaret M. Note on C. S. Peirce's Experimental Discussion of the Law of Errors. **Proceedings of the National Academy of Sciences of the United States of America**, National Academy of Sciences, v. 15, n. 2, p. 120–125, 1929. ISSN 00278424. Disponível em: <<http://www.jstor.org/stable/85603>>.

XIAO, Zhijie; KOENKER, Roger W. Conditional Quantile Estimation for Generalized Autoregressive Conditional Heteroscedasticity Models. **Journal of the American Statistical Association**, Taylor & Francis, v. 104, n. 488, p. 1696–1712, 2009. DOI: [10.1198/jasa.2009.tm09170](https://doi.org/10.1198/jasa.2009.tm09170). eprint: <https://doi.org/10.1198/jasa.2009.tm09170>. Disponível em: <<https://doi.org/10.1198/jasa.2009.tm09170>>.

ZIETZ, Joachim; ZIETZ, Emily Norman; SIRMANS, G. Stacy. Determinants of House Prices: A Quantile Regression Approach. **The Journal of Real Estate Finance and Economics**, v. 37, n. 4, p. 317–333, 2008. Disponível em: <<https://EconPapers.repec.org/RePEc:kap:jrefec:v:37:y:2008:i:4:p:317-333>>.