

# ANÁLISE DE MODELOS DE REGRESSÃO QUANTÍLICA OBTIDOS A PARTIR DE DADOS IMOBILIÁRIOS

## RESUMO

Segundo a NBR 14653-2 (2011), ~~a técnica mais utilizada~~ <sup>uma técnica recomendada</sup> quando se deseja estudar o comportamento de uma variável dependente em relação a outras que são responsáveis pela variabilidade observada nos preços de imóveis, é a análise de regressão, <sup>linear</sup> contudo, a própria norma prevê o uso de outras ~~diversas~~ metodologias para avaliação de um imóvel. Esse trabalho objetiva aplicar técnicas de regressão quantílica a dados imobiliários para obtenção de modelos para diversos quantis de distribuição, analisando seus parâmetros e comparando-os com o modelo de regressão linear preconizado pela norma de avaliação. É apresentado ainda o intervalo de confiança para cada parâmetro de alguns desses modelos. Esse estudo permitiu uma análise pormenorizada da influência das variáveis explicativas sobre o valor total dos imóveis em diversos quantis de distribuição.

**Palavras-Chave:** Regressão Quantílica; Mercado Imobiliário; Avaliação de Imóveis.

## 1 INTRODUÇÃO

Segundo Sheppard (1999), a determinação de preços de imóveis urbanos é objeto de extensa discussão teórica e pesquisa empírica já há algum tempo.

A engenharia de avaliações serve para subsidiar tomadas de decisões a respeito de valores, custos e alternativas de investimentos, envolvendo bens de qualquer natureza (DANTAS, 2012), como imóveis, sejam eles urbanos ou rurais.

Segundo Abunahman (2008, p. 10), uma considerável parcela de bens públicos e particulares consiste em bens imóveis. A própria amplitude desse recurso primordial em nossa sociedade cria uma necessidade de informes avaliatórios como suporte e consistência para decisões relativas ao uso e disposição desses bens.

Dentre os métodos mais utilizados para se determinar o valor de um bem, está o método comparativo direto de dados de mercado, utilizando modelos de regressão linear múltipla, amplamente aceito no meio científico. Nesse método, a estimação dos parâmetros do modelo se dá pelos mínimos quadrados, onde se busca minimizar a soma dos quadrados dos erros.

Segundo Dantas (2012), a regressão linear múltipla deve ser adotada quando mais de uma variável independente é necessária para explicar a variabilidade dos preços praticados no mercado. Com base em uma amostra extraída do mercado, os parâmetros populacionais são estimados por inferência estatística.

Esse estudo, no entanto, procura aplicar a dados imobiliários uma outra técnica, chamada de regressão quantílica, que objetiva obter modelos para os diversos quantis e que busca a minimização dos erros absolutos ponderados.

~~Concordando com Mosteller (1977, apud Koenker, 2005), o que a curva de regressão faz é dar um grande resumo das médias das distribuições correspondentes ao conjunto dos  $x$ 's observados. Poder-se-ia ir além e computar diversas curvas de regressão correspondentes aos diversos pontos percentuais da distribuição e dessa forma ter uma visão mais completa do conjunto.~~

Nesse sentido, esse estudo propõe fazer uma análise comparativa entre um modelo obtido por regressão linear múltipla, conforme preconiza a NBR 14.653-2, com modelos de regressão quantílica para a mediana ( $\tau = 0,50$ ), e para os quantis  $\tau = 0,10$ ;  $\tau = 0,20$ ,  $\tau = 0,30$ ,  $\tau = 0,40$ ,  $\tau = 0,60$ ,  $\tau = 0,70$ ,  $\tau = 0,80$ ,  $\tau = 0,90$ .

## 2 REGRESSÃO QUANTÍLICA

Os modelos de regressão são extremamente úteis em estudos estatísticos principalmente pela sua facilidade de interpretação. Segundo Santos (2012), dentre os métodos de estimação dos parâmetros do modelo, podemos citar o de minimização dos quadrados dos erros como o mais frequentemente utilizado.

Isso se deve, entre outros motivos, pelos bons resultados que se pode obter quando os erros do modelo apresentam distribuição normal.

Contudo, ainda que o método dos mínimos quadrados seja o mais utilizado, ele traz consigo algumas limitações, principalmente quando os erros do modelo não apresentam distribuição normal, exibindo uma distribuição assimétrica.

Essas limitações, que em geral ferem os pressupostos básicos do modelo de regressão, levaram os pesquisadores à busca por métodos diferenciados.

Segundo Koenker e Basset (1982), o que a curva de regressão pelo método dos mínimos quadrados faz é dar um grande resumo das médias das distribuições correspondentes ao conjunto dos  $x$ 's observados. Poder-se-ia ir mais longe e calcular diversas curvas de regressões diferentes correspondente aos vários pontos percentuais da distribuição e assim obter uma visão mais completa desse conjunto.



Um modelo de regressão usual ajustado pelo método dos mínimos quadrados ordinários resume toda informação das variáveis independentes observadas às suas médias. A média é uma informação resumida e incompleta de uma distribuição, da mesma forma que a regressão é uma visão limitada de um conjunto de distribuições. Uma possibilidade bem mais completa seria ajustar diversas curvas de regressão referentes a diversos quantis da distribuição (KOENKER, 2005).

Além disso, conforme Santos (2012), há a questão da influência que *outliers* exercem nas estimativas dos parâmetros do modelo. Isso faz com que seja necessário sempre que se utiliza essa técnica, uma criteriosa avaliação de quanto cada ponto influencia no ajuste do modelo, o que se torna bastante trabalhoso.

Nessa linha, Santos (2012) nos mostra que o método de minimização dos erros absolutos é robusto na presença de *outliers* na variável resposta.



Quando a distribuição dos erros não é normal, esse método se mostra melhor para descrever uma posição central da distribuição condicional da variável resposta, ao estimar o valor mediano da distribuição. Já a regressão quantílica se baseia no método dos erros absolutos, porém, para estimar os diversos quantis de interesse, é feita uma ponderação na minimização desses erros (SANTOS, 2012).

A teoria relacionada à regressão quantílica foi proposta por Koenker e Basset (1978) e permite calcular estimativas para quaisquer quantis da variável resposta, tais como a mediana. Isso permite com que se obtenha mais informações de localização, como na cauda inferior ou superior da distribuição.



Desta forma, é possível avaliar o impacto de uma covariável em toda a distribuição da variável resposta, e não apenas na sua média. Segundo Silva *et al* (2006), os diferentes resultados de uma variável explicativa em quantis distintos podem ser interpretados como diferenças na resposta da variável dependente a mudanças nos regressores em vários pontos na distribuição condicional.

Segundo Araújo Júnior (2018, p. 8), esses efeitos quantílicos indicam que os atributos podem impactar diferentemente a variável de resposta de acordo com o quantil de sua distribuição condicional.

Essa classe de modelos de regressão se trata de uma abordagem mais geral que os modelos clássicos e podem ser caracterizados como modelos de regressão semiparamétricos, pois não exigem nenhuma distribuição de probabilidade para a variável resposta (SOUZA, 2017).

Do mesmo modo, como citado por Santos (2012), os modelos de regressão quantílica são robustos em relação a *outliers* e erros que não seguem a distribuição normal, podendo incorporar uma possível heterocedasticidade detectada a partir da variação das estimativas dos coeficientes dos parâmetros para os diferentes quantis da distribuição (BARROSO *et al.*, 2015).

Conforme Koenker (2005), o quantil de ordem  $\tau$  de uma variável aleatória  $Y$  pode ser definido como a função inversa de distribuição acumulada no ponto  $\tau$ :

$$Q_{\tau}(Y) = F^{-1}(\tau) = \inf \{y \mid F(y) \geq \tau\} \quad (1.1)$$

Com  $F(y) = P(Y \leq y)$  a função de distribuição acumulada de  $Y$ , com  $\tau \in [0, 1]$ . Quando  $\tau = 0,50$ , temos  $F^{-1}(0,50)$ , caso em que é estimada a mediana condicional. O primeiro quartil e o terceiro quartil são  $F^{-1}(0,25)$  e  $F^{-1}(0,75)$ , respectivamente.

Considerando então uma amostra obtida de uma população cujos pares ordenados são  $(y_i, x_i)$  com  $i = 1, 2, 3, \dots, n$ , para cada  $x$ , é possível relacionar a variável resposta  $y$  e o vetor de covariáveis  $x$ , por meio de uma relação linear do tipo:

$$Q_{\tau}(y_i \mid x_i) = y_i = \beta_0(\tau) + \beta_1(\tau)x_{i1} + \dots + \beta_p(\tau)x_{ip} \quad (1.2)$$

Em que  $Q_{\tau}(y_i \mid x_i)$  é o quantil condicional  $y_i \mid x_i$  e os valores  $\beta_0(\tau), \dots, \beta_p(\tau)$  são os parâmetros desconhecidos indexados no quantil  $\tau$ .

Pode-se definir os quantis de uma amostra da seguinte forma: “O quantil de ordem  $\tau$  de uma amostra é o valor  $m$  tal que  $100\tau\%$  dos valores amostrais são inferiores a ele, com  $0 < \tau < 1$  (SANTOS, 2012).

Koenker e Machado (1999) mostram que é possível estimar os parâmetros desconhecidos, indexados no quantil  $\tau$ , diretamente através da solução do problema de minimização definido por:

$$\min_{\beta \in \mathbb{R}^p} \sum_{i=1}^n \rho_{\tau} \cdot [y_i - (\beta_0(\tau) + \beta_1(\tau)x_{i1} + \dots + \beta_p(\tau)x_{ip})] \quad (1.3)$$



Em que  $\rho(\tau)$  é um termo ponderador. Desta forma, o  $\tau$ -ésimo estimador  $\hat{\beta}(\tau)$  da regressão quantílica é selecionado de modo a minimizar a soma dos desvios do  $y$  observado até um valor de  $\hat{y}$  ajustado.

Segundo Nascimento *et al* (2012), o modelo de regressão quantílica permite ao investigador ter uma visão mais completa da relação existente entre a variável resposta e as covariáveis observadas, uma vez que é possível construir um modelo para cada quantil de interesse.

### 2.2.1 Qualidade do 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:

$$Q_\tau(y_i | x_i) = x_i' \beta(\tau) = x_{i1}' \beta_1(\tau) + x_{i2}' \beta_2(\tau) \quad (1.4)$$

Em que  $x_i$ ,  $i$ -ésima linha da matriz de planejamento  $X$ , é particionada em duas partes aqui chamadas de  $x_{i1}$  e  $x_{i2}$  de dimensão  $p - q$  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:

$$\hat{V}(\tau) = \sum_{i=1}^n \rho_\tau \cdot [y_i - x_i' \hat{\beta}(\tau)] \quad (1.5)$$

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:

$$\tilde{V}(\tau) = \sum_{i=1}^n \rho_\tau \cdot [y_i - x_{i1}' \tilde{\beta}_1(\tau)] \quad (1.6)$$

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

$$R^1(\tau) = 1 - \frac{\hat{V}(\tau)}{\tilde{V}(\tau)} \quad (1.7)$$

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  $R^2$  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 uma apropriada 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 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).

### 3 ESTUDO DE CASO

O presente estudo tem caráter descritivo com uma abordagem quantitativa e inferencial, onde se busca fazer uma análise de modelos gerados por regressão quantílica, ajustados aos diferentes quantis de interesse.

A amostra utilizada no estudo é um conjunto com 50 dados de mercado do centro de Florianópolis coletados em 2015 (Anexo A). Os dados possuem como variável explicada o valor **unitário (VU)** dos imóveis, em **R\$/m<sup>2</sup>** e como variáveis explicativas a área total do imóvel coletado em m<sup>2</sup>, o números de quartos, o número de suítes, o número de vagas de garagem, a distância à beira mar em metros, e o padrão construtivo do imóvel, totalizando 6 variáveis explicativas.



**Figura 01:** distribuição espacial dos dados de mercado da amostra.



Todos os 50 dados coletados foram tabulados em uma planilha eletrônica e na sequência transferidos para o *software* R na sua versão 3.5.1, plataforma livre com diversos métodos estatísticos de análise de dados já implementados.

Para se gerar os modelos de regressão utilizou-se o pacote *quantreg* (KOENKER, 2005), disponível no *software* R.

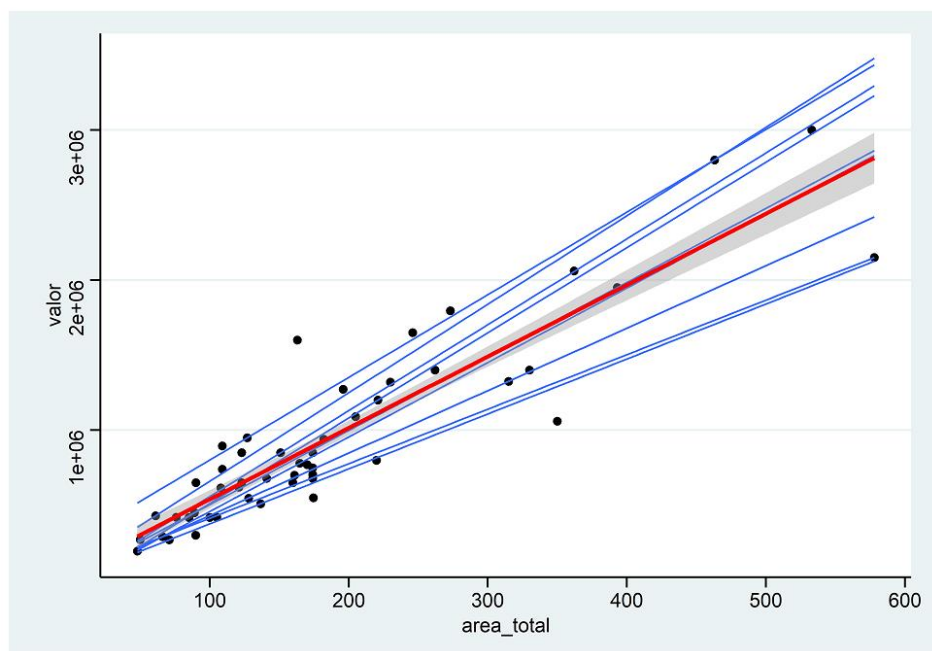
Na sequência, fez-se uma análise dos resultados das estatísticas obtidas com os dados de mercado para os diversos modelos de regressão gerados.

### 3.1 CASO EM DUAS DIMENSÕES

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

Seja primeiramente o caso de dados heteroscedásticos. A figura 2 ilustra a aplicação da regressão quantílica e da regressão linear para este caso. A linha 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 linhas azuis são retas de regressão quantílica para quantis 0,1; 0,2; 0,3; 0,4; 0,5; 0,6; 0,7; 0,8 e 0,9.

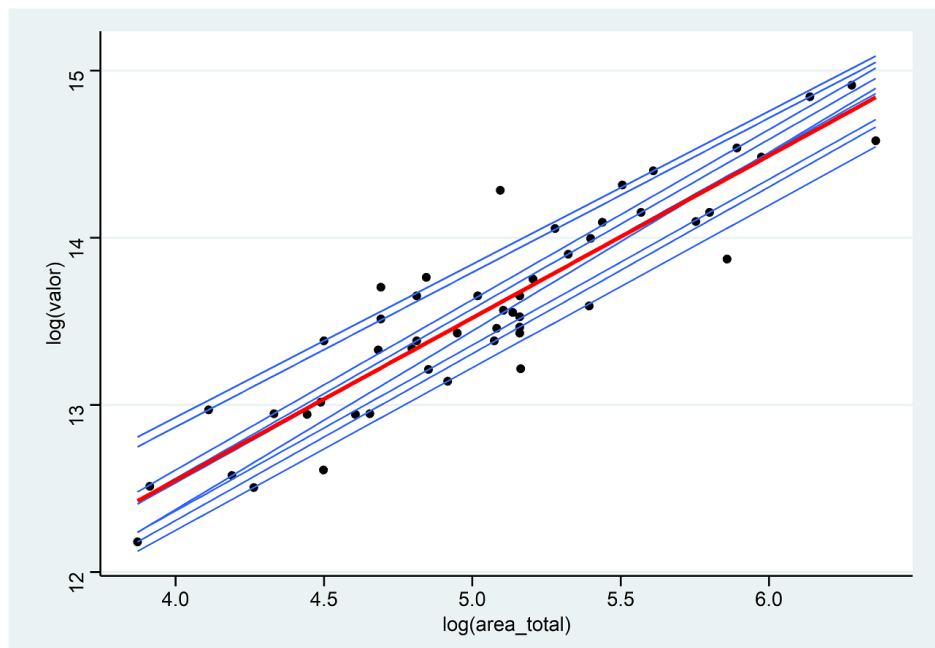
A regressão quantílica neste caso pode ser usada para demonstrar a não validade dos intervalos de confiança (*IC*) e predição (*IP*) para a regressão linear para este tipo de dado: como a variância da população não é constante, mas aumenta com o aumento da área, as retas da regressão quantílica se abrem. 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 intervalo de confiança *IC*.



**Figura 02:** Regressão Linear e Quantílica para dados heteroscedásticos.

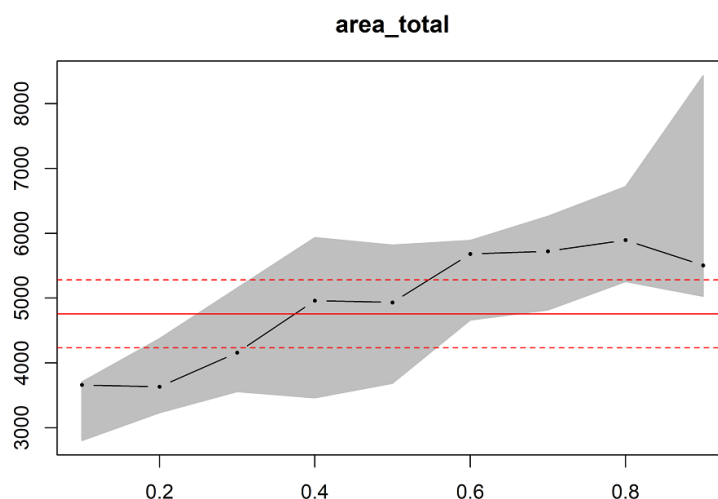
Assim como na regressão linear, uma conveniente transformação das variáveis pode ser aplicada para a obtenção da homocedasticidade. Isto pode ser

visto na figura 3 abaixo, onde as retas para os diferentes quantis obtidas pela regressão quantílica agora são praticamente paralelas entre si, indicando que a heteroscedasticidade foi removida.



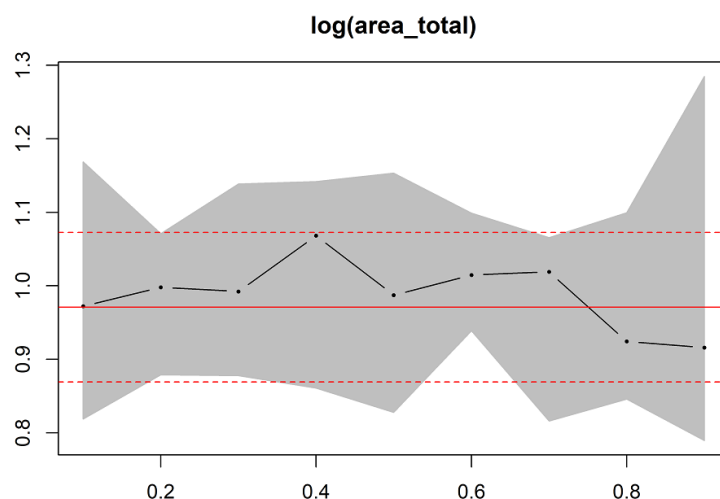
**Figura 03:** Regressão Linear e Quantílica com dados transformados.

Os coeficientes das retas de regressão quantílica podem ser plotados como na figura 4. 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 significativamente diferentes, estatisticamente, do coeficiente da regressão linear.



**Figura 04:** variação dos coeficientes de regressão quantílica (variáveis originais).

Já para os dados transformados, pode-se notar na figura 5 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 (MATLOFF, 2017).

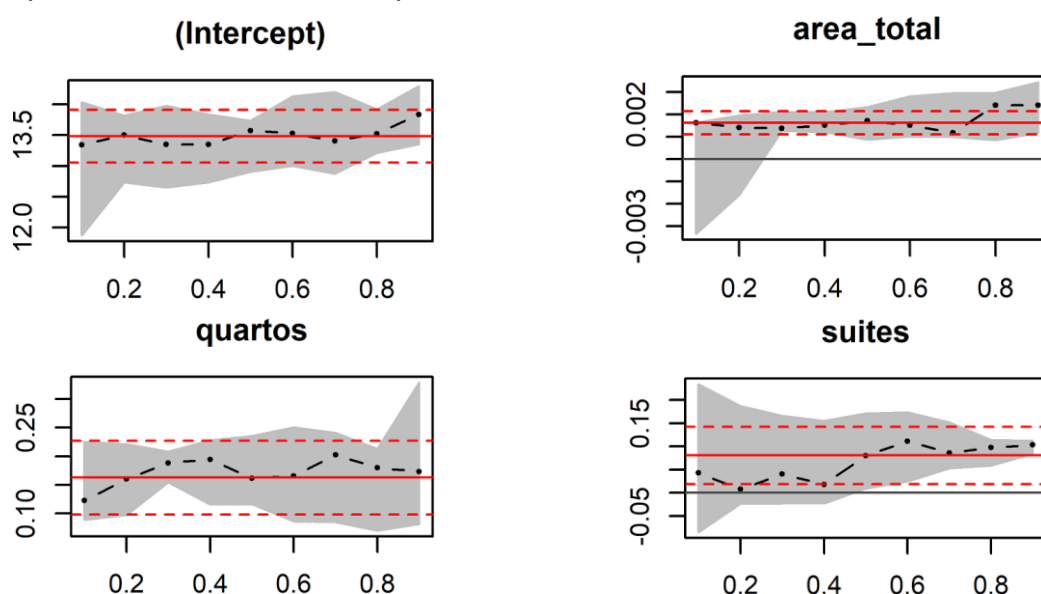


**Figura 05:** variação dos coeficientes de regressão quantílica (variáveis transformadas).

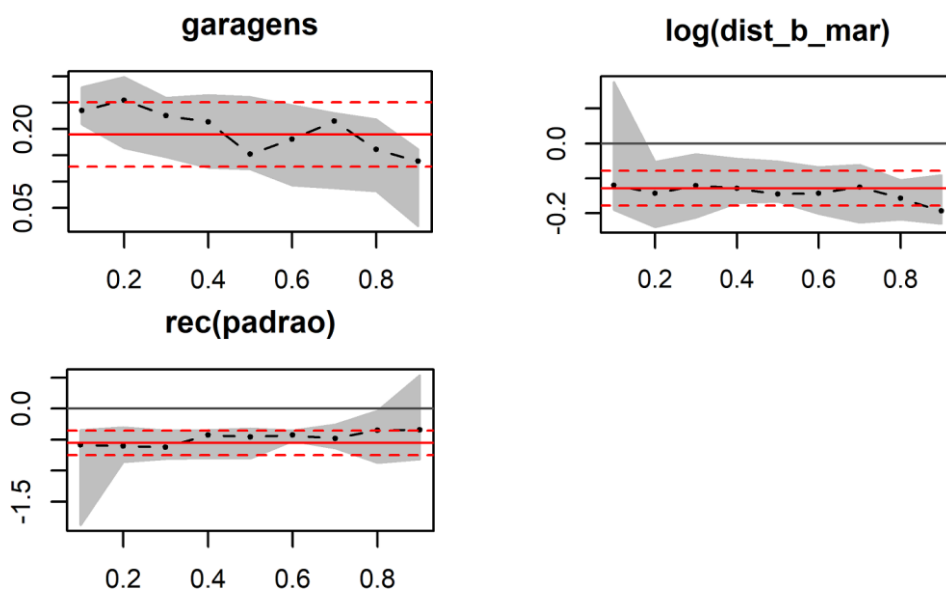
### 3.2 ANÁLISE MULTIVARIADA

Para a amostra (Anexo A) 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 quantis 0,1; 0,2; 0,3; 0,4; 0,5; 0,6; 0,7; 0,8 e 0,9.

Na figura 6 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 significativamente dos coeficientes da regressão linear, com exceção para alguns quantis superiores nas variáveis área total e padrão construtivo, como pode ser visto abaixo:







**Figura 06:** coeficientes de regressão linear e quantílica na análise multivariada.

Na tabela 1 podem ser vistos os coeficientes e estatísticas básicas dos modelos de regressão linear e de regressão quantílica à mediana ( $\tau = 0,5$ ).

VARIÁVEL DEPENDENTE: LOG (VALOR_TOTAL)		
	Regressão Linear (1)	Regressão Quantílica (2)
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 ***
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 ***

Observations	48	50
R <sup>2</sup>	0.956	
Adjusted R <sup>2</sup>	0.950	
Res. Std. Erro	0.136 (df = 41)	
F Statistic	148.921 *** (df = 6; 41)	

Nota: \*p < 0.1; \*\*p < 0.05; \*\*\*p < 0.01

**Tabela 01:** comparação entre os modelos de regressão linear e regressão à mediana.

### 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 estimação 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 7,83%.

de 80%



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 13,10%.

O modelo de regressão linear mostrou-se, portanto, mais eficiente do que o modelo de regressão à mediana, apesar do 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.

Para o modelo de regressão quantílica, o intervalo de predição *IP* não faz qualquer sentido. No entanto, é possível estimar os valores diretamente para os quantis 0,1 e 0,9 da população. Neste 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 quantis 0,1 e 0,9. Os limites inferior e superior do intervalo de confiança *IC* para o quantil 0,1 são, respectivamente: R\$ 781.253,06 e R\$ 841.110,17. Os limites inferior e superior do intervalo de confiança *IC* para o quantil 0,9 são, respectivamente: R\$ 1.116.547,53 e R\$ 1.261.800,41.

## 5 CONSIDERAÇÕES FINAIS

Não se pode negar que a análise de regressão é uma das técnicas estatísticas mais adotadas para se realizar a análise de dados. Contudo, novas técnicas que busquem melhorar o entendimento da relação existente entre a variável resposta e as variáveis explicativas devem ser estudadas e difundidas.


Os resultados demonstraram que as características de um imóvel, como área total, distância a um ponto de valorização e número de suítes ou vagas de garagem, tem influências diferentes sobre os valores totais ao longo dos diversos quantis.

valorização  
valorização



Nesse estudo, os modelos gerados por regressão quantílica permitiram avaliar a relação das variáveis coletadas além da posição central correspondente à média ou a mediana. Sendo assim, pode-se avaliar o efeito das diversas variáveis

sobre o valor total dos imóveis em diversos pontos da distribuição condicional dessas variáveis, como na cauda inferior relativa ao quantil da ordem de 0,10 ou também na cauda superior relativa ao quantil da ordem de 0,90.

Verificou-se com esse estudo que as estimativas realizadas com os modelos de regressão quantílica para os quantis 0,10 e 0,90 foram semelhantes com as estimativas dos limites do intervalo de predição da regressão linear, ou seja, constatou-se que o intervalo de predição <sup>levou</sup> ~~leva~~ aos valores extremos da população. 

Salienta-se a possibilidade de se calcular outros quantis com a regressão quantílica, sem que haja necessidade de correção da heterocedasticidade.

Observou-se ainda que a regressão à média é mais eficiente que a regressão à mediana pois, mesmo com dois dados a menos (4% do total de dados), a regressão à média teve intervalo de confiança de amplitude com aproximadamente metade do intervalo de confiança da regressão à mediana.

Desta forma, é importante destacar que, em um contexto de *Big Data*, a regressão quantílica à mediana pode ser uma alternativa interessante, haja vista que ela é robusta a presença de *outliers*, não necessitando de análises mais detalhadas.

*A regressão quantílica também surge como opção em estudos de mercado quando a heterocedasticidade não pode ser eliminada na construção de modelos de regressão lineares.*

## REFERÊNCIAS BIBLIOGRÁFICAS

ABNT. ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **NBR 14653-2:** Avaliação de Bens. Parte 2: Imóveis Urbanos. Rio de Janeiro, 2011.

ABUNAHMAN, Sérgio Antônio. **Curso básico de engenharia legal e de avaliações**. 4. ed. São Paulo: Pini, 2008.

ARAÚJO JÚNIOR, Ignácio Tavares de. **O uso de regressões quantílicas na avaliação de imóveis**. In: Simpósio da Sociedade Brasileira de Engenharia de Avaliações; nov. 2018; João Pessoa: SOBREA, 2018.

BARROSO, L. M. A.; NASCIMENTO, M.; NASCIMENTO, A. C. C., SILVA, F. F. e; CRUZ, C. D.; BHERING, L. L.; FERREIRA, R. de P. **Metodologia para análise de adaptabilidade e estabilidade por meio de regressão quantílica**. Pesquisa Agropecuária Brasileira, Brasília, v. 50, n. 4, p. 290-297. 2015.

DANTAS, Rubens Alves. **Engenharia de avaliações:** uma introdução à metodologia científica. 3. ed. São Paulo: Pini, 2012. 255 p.

KOENKER, R. **Quantile regression**. 1. ed. Cambridge University Press, v.1. 349 p, 2005.

KOENKER R. **Quantreg: quantile regression**. R package version 5.21, 2016. Disponível em: <<https://cran.r-project.org/web/packages/quantreg/index.html>>. Acesso em: 15/07/2019.

KOENKER, R.; BASSETT, G. **Regression quantiles**. Econometrica, v. 46, 1978.

KOENKER, R.; PARK, B. J. **An interior point algorithm for nonlinear quantile regression**. Journal of Econometrics, 71, 265-283. 1996.

KOENKER, R.; MACHADO, J. A. F. **Goodness of fit and related inference processes for quantile regression.** Journal of the American Statistical Association, v. 94, n. 448, p.1296-1310, 1999.

MATLOFF, N. **From linear models to machine learning:** regression and classification, with R examples. Chapman & Hall, 2017.

SANTOS, Bruno Ramos. **Modelos de regressão quantílica.** Dissertação (Mestrado em Ciências) - Programa de Pós-Graduação em Matemática e Estatística, USP, São Paulo, 2012.

SHEPPARD, S. **Hedonic analysis of housing markets.** In: Handbook of regional and urban economics. Amsterdam: North-Holland, 1999. p.1595-1636.

SILVA, E. N. da; PORTO JÚNIOR, S. da S. **Sistema financeiro e crescimento econômico: uma aplicação de regressão quantílica.** Economia Aplicada, Ribeirão Preto, v. 10, n. 3, p. 425-442.

SOUZA, Saul de Azevedo. **Modelagem da obesidade adulta nas nações: uma análise via modelos de regressão quantílica e beta.** Dissertação (Mestrado em Ciências Exatas e da Natureza) - Programa de Pós-Graduação em Modelos de Decisão e Saúde, UFPB, Paraíba, 2017.

## ANEXO A – DADOS DE MERCADO UTILIZADOS NO ESTUDO

IDENT	Valor_Total	Area_Total	Quartos	Suítes	Garagens	Dist_B Mar	Padrão
AP_01	1.060.000,00	350,00	3	1	2	720	2
AP_02	510.000,00	136,56	3	1	1	665	2
AP_03	780.000,00	164,77	3	1	2	415	2
AP_04	550.000,00	174,58	3	1	1	320	2
AP_05	850.000,00	123,01	3	1	3	895	3
AP_06	300.000,00	89,83	2	0	1	645	1
AP_07	750.000,00	174,00	2	1	2	860	3
AP_08	650.000,00	123,00	3	1	1	745	3
AP_09	620.000,00	121,00	3	1	1	745	3
AP_10	740.000,00	109,00	3	1	1	300	2
AP_11	770.000,00	170,00	3	1	2	590	2
AP_12	680.000,00	141,00	3	1	1	290	2
AP_13	850.000,00	174,00	3	1	1	465	2
AP_14	420.000,00	105,00	3	1	0	60	1
AP_15	547.000,00	128,00	3	1	1	745	3
AP_16	1.600.000,00	163,00	4	2	2	90	3
AP_17	1.320.000,00	230,00	3	1	2	215	3
AP_18	615.000,00	108,00	3	1	1	745	3
AP_19	705.000,00	174,00	2	1	2	900	3
AP_20	418.000,00	85,00	1	0	1	620	3
AP_21	270.000,00	71,00	2	0	0	1380	1
AP_22	418.000,00	100,00	1	1	1	620	3
AP_23	650.000,00	90,00	2	1	1	215	3
AP_24	700.000,00	161,00	2	1	2	500	3
AP_25	680.000,00	174,00	2	1	2	860	3
AP_26	420.000,00	76,00	2	1	1	700	1
AP_27	195.000,00	48,00	1	0	0	730	1
AP_28	290.000,00	66,00	1	0	1	745	1
AP_29	272.000,00	50,00	1	0	1	1430	1
AP_30	430.000,00	61,00	2	0	1	170	1
AP_31	895.000,00	109,00	3	1	1	530	2
AP_32	450.000,00	89,00	2	0	1	745	2
AP_33	1.950.000,00	393	3	1	3	550	3
AP_34	2.150.000,00	578	3	2	3	260	3
AP_35	940.000,00	182	3	1	2	200	2
AP_36	1.400.000,00	262	4	1	1	60	3
AP_37	1.090.000,00	205,00	3	0	3	465	2
AP_38	1.272.000,00	196	3	3	2	610	3
AP_39	2.800.000,00	463	3	3	3	590	3
AP_40	1.796.000,00	273	3	3	4	140	2
AP_41	1.400.000,00	330	4	2	2	655	3
AP_42	3.000.000,00	533	4	3	4	427	3
AP_43	1.200.000,00	221	3	3	2	607	3
AP_44	800.000,00	220	3	1	1	1000	2
AP_45	950.000,00	127	2	1	1	60	2
AP_46	2.061.000,00	362,00	3	3	4	310	3
AP_47	1.326.000,00	315,00	3	3	3	600	3
AP_48	850.000,00	151,00	3	1	2	660	2
AP_49	1.650.000,00	246,00	3	3	3	307	3
AP_50	650.000,00	159,72	3	1	1	120	2