

MODELO DE REGRESSÃO LINEAR MÚLTIPLA PARA AVALIAÇÃO DE IMÓVEIS RESIDENCIAIS EM FRANCISCO BELTRÃO, PARANÁ

GUSTAVO ADRIANO BOHN^{1*}, LILIA ANDREIA CARDOSO BOHN²

¹Engenheiro Civil, UNIOESTE, Cascavel-PR, Rua Padre Cirilo, 150, Centro, Capanema/PR, CEP 85.760-000, gustavo.bohn@gmail.com

²Discente de Engenharia Ambiental, UNISEP, Dois Vizinhos-PR, lilia.bohn@gmail.com

Apresentado no
Congresso Técnico Científico da Engenharia e da Agronomia – CONTECC'2016
29 de agosto a 1 de setembro de 2016 – Foz do Iguaçu, Brasil

RESUMO: Este artigo tem por objetivo encontrar um modelo de regressão linear com múltiplas variáveis para avaliação dos imóveis residenciais da cidade de Francisco Beltrão, Estado do Paraná. Coletaram-se trinta dados de operações imobiliárias residenciais para compor a amostra, as características analisadas incluem a área do terreno, a localização, a área da obra e o padrão de acabamento das residências; transformadas em seis variáveis explicativas que foram processadas e interpretadas quanto ao comportamento na formação do preço. O método utilizado foi o Comparativo Direto de Dados de Mercado, através do tratamento científico que emprega técnicas de inferência estatística. Após várias simulações em busca do melhor ajustamento de dados, o coeficiente de correlação obtido foi de 89,14%, indicando uma forte correlação entre a variável dependente e as independentes. Pela análise de sensibilidade, a equação gerada demonstrou-se adequada à realidade do mercado imobiliário residencial da cidade.

PALAVRAS-CHAVE: Avaliação. Imóveis residenciais. Inferência estatística.

MULTIPLE LINEAR REGRESSION MODEL FOR EVALUATION OF RESIDENTIAL PROPERTIES IN FRANCISCO BELTRÃO, PARANÁ

ABSTRACT: This article aims to find a linear regression model with multiple variables for evaluation of residential properties in the city at Francisco Beltrão, State of Paraná. It was collected thirty data of residential real estate transactions for the sample, the characteristics analyzed include the land area, the location, the construction area and the residential standard of finish; transformed into six explanatory variables that were processed and interpreted for price formation. The method used was the Direct Comparison of Market Data, through scientific treatment employing techniques of statistical inference. After several simulations to find the best fit data, the correlation coefficient obtained was 89.14%, indicating a strong correlation between the dependent and independent variables. For sensitivity analysis, the equation generated demonstrated to be appropriate to the reality of the residential real estate market of the city.

KEYWORDS: Evaluation. residential real estate. Statistical inference.

INTRODUÇÃO

A Engenharia de Avaliações é uma especialidade da engenharia que reúne um conjunto amplo de conhecimentos com o objetivo de determinar tecnicamente o valor de um bem, de seus direitos, frutos e custos de reprodução (Dantas, 2012). De uma maneira geral, quando o bem a ser avaliado for um imóvel, busca-se identificar essencialmente o valor financeiro frente ao mercado onde ele é transacionado. Conforme a NBR 14653-1 (2001), que trata dos procedimentos gerais para avaliação de bens, o Valor de Mercado é a “quantia mais provável pela qual se negociaria voluntariamente e conscientemente um bem, numa data de referência, dentro das condições do mercado vigente”.

A escolha da metodologia mais adequada para realização de um trabalho avaliatório depende fundamentalmente das condições mercadológicas com que se defronta o avaliador, pelas informações coletadas neste mercado, bem como a natureza do serviço que se pretende desenvolver. Contudo,

“para a identificação do valor de mercado, sempre que possível preferir o método comparativo direto de dados de mercado” (NBR 14653-1, 2001). Neste método, o valor do imóvel será obtido pela comparação de dados de mercado relativos a outros de características similares, através do tratamento técnico de seus atributos (Fiker, 2008).

O método comparativo direto de dados de mercado nos possibilita a utilização de dois tipos de tratamentos, o tratamento por fatores e o tratamento científico. Tendo em vista que no tratamento científico a estimativa do valor é realizada utilizando-se modelos elaborados especificamente para avaliação do bem avaliando, pela substituição de suas características na equação resultante, pode-se imprimir maior nível de precisão e fundamentação ao trabalho (Dantas, 2012).

Sendo assim, para o desenvolvimento do presente trabalho, optou-se pelo tratamento científico com o uso da estatística inferencial, que permite explicar o valor de mercado através das diversas variáveis que influenciam na sua formação. Adotando-se esse método, buscou-se construir um modelo de regressão linear múltipla para avaliação de imóveis residenciais localizados na cidade de Francisco Beltrão, Estado do Paraná, a partir de uma amostra de dados coletada. Este trabalho encerra com uma análise criteriosa dos resultados obtidos. Com isso, propõe-se a oferecer uma contribuição à área de Engenharia de Avaliações, mais especificamente na avaliação de imóveis urbanos da tipologia residencial.

MATERIAIS E MÉTODOS

Inicialmente, procedeu-se o levantamento de dados de mercado para compor a amostra dos imóveis residenciais da cidade de Francisco Beltrão, transacionados em época contemporânea ao estudo. Obteve-se um banco de dados contendo 30 dados de amostra de imóveis residenciais descritos por seus atributos, em vista da formação das variáveis independentes e da variável dependente. As características consideradas na construção do modelo são as seguintes:

↯ Área do terreno: Refere-se à área do terreno do imóvel, em metros quadrados;

↯ Localização: Separou-se esta característica em Localização Ruim, Localização Regular e Localização Boa, conforme o padrão usual de acabamento do imóvel pesquisado. O padrão usual de acabamento foi definido como sendo a Moda entre os padrões de acabamento dos imóveis internos à região circular de raio 200m em torno deste imóvel. Para caracterização do padrão de acabamento de cada imóvel interno a região considerada de influência direta ao imóvel pesquisado, utilizou-se a classificação proposta por Fiker (2008). A Localização Ruim está associada aos imóveis com padrão usual de acabamento simples, enquanto a Localização Regular engloba os imóveis com padrão usual de acabamento médio e, finalmente, a Localização Boa abrange os imóveis com padrão usual de acabamento superior.

↯ Área da obra: Refere-se à área da edificação residencial presente no imóvel. Neste caso, considerou-se a área equivalente da construção ao invés da área real, calculada de acordo com o previsto na NBR 12721 (2006) a fim de ajustar os dados que apresentam área de padrão diferente do paradigma;

↯ Padrão de acabamento: As residências foram separadas em Padrão Simples, Padrão Médio e Padrão Superior, quanto ao padrão de acabamento, de acordo com a classificação proposta por Fiker (2008).

Em seguida, foram codificadas as variáveis qualitativas para que a planilha apresentasse apenas dados numéricos, possibilitando a análise dos dados nela contidos. As variáveis “localização” e “padrão de acabamento” foram convertidas em dicotômicas de grupo, formando quatro variáveis: a localização boa (LB), a localização regular (LR), o padrão de acabamento superior (PS) e o padrão de acabamento médio (PM), que agregadas a área do terreno (AT) e a área da obra (AO) formaram o grupo das variáveis explicativas, composto por seis variáveis. Com relação a variável dependente ou explicada, utilizou-se o valor unitário do imóvel (VU), que representa o quociente do valor do imóvel pela área da obra.

Para a construção do modelo foi utilizado o programa computacional TS-Sisreg® (Sistema de Regressão Linear Múltipla), versão 1.5.3, desenvolvido pela Tecsys Engenharia. O programa gerou vários modelos, dentre os quais, procurou-se a melhor equação de regressão que representa o comportamento dos valores de mercado.

RESULTADOS E DISCUSSÕES

A equação de regressão que melhor ajustou os valores coletados (também chamados de valores observados) aos valores calculados apresentou um coeficiente de correlação (r) no valor de 0,8914 ou 89,14%, isso significa uma forte correlação entre a variável dependente e as variáveis independentes utilizadas no modelo. O coeficiente de determinação (R) encontrado foi de 0,7946, o que significa dizer que 79,46% da variabilidade dos valores observados são devido as variáveis adotadas enquanto que 20,54% dos valores observados não foram explicados pelo modelo de regressão por prováveis variáveis não consideradas ou erros ocasionais de medidas.

Para verificação da significância individual de cada regressor (variável independente), utilizaram-se os resultados da Tabela 1, onde se encontra o valor de T-Student para cada variável utilizada no modelo com a sua respectiva significância. O modelo gerado apresentou Grau III para este item da Norma, já que os resultados da significância ficaram entre 0,01% e 4,69%.

Tabela 1. T-Student e significância dos regressores.

VARIÁVEL	Escala Linear	T-Student Calculado	Significância % (Soma das Caudas)
AT	X	2,1004	4,69
LB	X	4,83531	0,01
LR	X	2,77946	1,07
AO	ln(x)	-6,44578	0,01
PS	X	6,44493	0,01
PM	X	2,78593	1,05

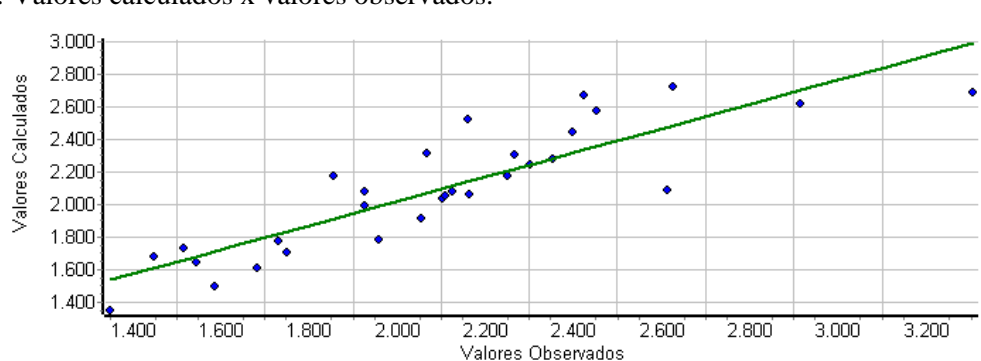
Fonte: TS-Sisreg®.

Na sequência, efetuou-se a análise da significância global do modelo, através do teste F. O F calculado na equação para uma significância de 1% foi de 14,83, já o valor de F tabelado para 6 graus de liberdade no numerador e 23 graus de liberdade no denominador é 3,71, como $F_c > F_{tab}$ ($14,83 > 3,71$) rejeita-se H_0 , o que é desejável. Como o F calculado para a significância de 1% atende a rejeição da hipótese nula, então o modelo recebe o Grau III de fundamentação para este item da Norma.

Já na análise de normalidade, os percentuais dos resíduos padronizados devem apresentar uma tendência à distribuição normal. Examinando o histograma dos resíduos amostrais padronizados, verificou-se que sua forma guarda semelhança com a curva normal. As frequências relativas dos resíduos amostrais padronizados nos intervalos de $[-1; +1]$, $[-1,64; +1,64]$ e $[-1,96; +1,96]$ apresentaram-se próximas às probabilidades da distribuição normal padrão estabelecidas pela NBR 14653-2 (2011), desta forma pode-se assegurar a normalidade dos resíduos com os seguintes intervalos: 73% dos resíduos padronizados estão no intervalo $[-1; +1]$; 93% dos resíduos padronizados estão no intervalo $[-1,64; +1,64]$ e 96% dos resíduos padronizados estão no intervalo $[-1,96; +1,96]$.

A Figura 1 ilustra o poder de predição do modelo que, segundo a NBR 14653-2 (2011), “deve ser verificado a partir do gráfico de valores observados na abscissa *versus* valores estimados pelo modelo na ordenada, que deve apresentar pontos próximos da bissetriz do primeiro quadrante”, ou seja, quanto mais os pontos se aproximarem da bissetriz (reta verde) maior será o poder de predição do modelo. Os resultados demonstram um alto poder de predição para a linearização escolhida.

Figura 1. Valores calculados x valores observados.



Fonte: TS-Sisreg®

A existência de pontos atípicos foi verificada pelo gráfico dos resíduos padronizados *versus* valores calculados, onde se constatou a presença de um elemento discrepante. Optou-se por manter na amostra este dado pesquisado, pois seu resíduo padronizado está abaixo de +2,5, representando 3,33% da amostra, enquanto que 96,67% dos resíduos permanecem dentro do intervalo -2 e +2.

Com relação à multicolinearidade, analisaram-se as correlações entre as variáveis independentes que são mostradas na matriz de correlação ilustrada na Tabela 2, onde na parte superior estão as correlações parciais e na parte inferior as correlações isoladas. A maior correlação foi de 0,76 (76%) entre as variáveis PS e AO, embora seja um valor elevado, enquadra-se à Norma 14.653-2 (2011), que recomenda atenção especial para resultados superiores a 0,80.

Tabela 2. Matriz de correlações.

variável	Forma linear	AT	LB	LR	AO	PS	PM	VU
AT	x		20	15	57	15	18	40
LB	x	30		68	65	59	29	71
LR	x	-7	-52		47	45	21	50
AO	ln(x)	68	32	-5		76	41	80
PS	x	60	20	-11	60		57	80
PM	x	-16	2	4	-17	-36		50
VU	ln(y)	30	40	-4	-7	47	11	

Fonte: TS-Sisreg®.

Concluída a análise dos pressupostos do modelo, obteve-se a equação de regressão linear múltipla para avaliação de imóveis residenciais na cidade de Francisco Beltrão, conforme Figura 2.

Figura 2. Equação de regressão linear múltipla.

$$VU = 8915,8736 * e^{(0,00038653653 * AT)} * e^{(0,27903189 * LB)} * e^{(0,12834671 * LR)} * AO^{-0,39261899} * e^{(0,39652279 * PS)} * e^{(0,13207952 * PM)}$$

Fonte: TS-Sisreg®

Na sequência, realizou-se a análise da sensibilidade do modelo de regressão para algumas residências, escolhendo-se aleatoriamente cinco dados da amostra, por representarem imóveis transacionados na mesma época e condição da pesquisa. A variação percentual entre os valores unitários observados na pesquisa e os valores unitários calculados pela equação pode ser verificada na Tabela 3. O modelo apresentou uma variação entre -7,16% e +3,52% em torno da estimativa de tendência central, permanecendo dentro do limite de amplitude de 15%, para mais e para menos, do campo de arbítrio estabelecido na NBR 14653-2 (2011).

Tabela 3. Variação entre valor coletado e calculado.

Dado da pesquisa (n°)	VU observado (R\$/m²)	VU calculado (R\$/m²)	Variação (%)
2	1.588,24	1.492,05	-6,05639
6	1.683,53	1.605,38	-4,64203
16	2.250,61	2.170,30	-3,56837
19	2.625,00	2.717,36	3,51848
30	2.057,14	1.909,80	-7,16237

Para a análise de sensibilidade quanto ao grau de precisão na estimativa dos valores unitários, o modelo de regressão gerado atendeu o Grau III, pois a amplitude do intervalo de confiança de 80% em torno da estimativa de tendência central permaneceu dentro do limite de 30% preconizado pela NBR 14653-2 (2011) para todos os imóveis, conforme Tabela 4.

Tabela 4. Amplitude do intervalo de confiança.

Dado da pesquisa (n°)	VU mínimo calculado (R\$/m ²)	VU médio calculado (R\$/m ²)	VU máximo calculado (R\$/m ²)	Amplitude (%)
2	1.404,19	1.492,05	1.585,41	12,15
6	1.512,48	1.605,38	1.703,98	11,93
16	2.050,59	2.170,30	2.297,00	11,35
19	2.448,09	2.717,36	3.016,24	20,91
30	1.802,60	1.909,80	2.023,38	11,56

Para o enquadramento quanto ao grau de fundamentação, verificou-se que o modelo atende há três itens do Grau III e três itens do Grau II. Contabilizaram-se três pontos para cada item no Grau III e dois pontos para cada item do Grau II de acordo com a orientação da NBR 14653-2 (2011), totalizando 15 pontos, e comparou-se com a tabela de enquadramento para modelos de regressão linear, o que enquadra o modelo de regressão no Grau II de fundamentação.

CONCLUSÃO

Foi possível obter um resultado satisfatório para o modelo, muito embora tenha se utilizado um número reduzido de dados coletados. Isso se deve a qualidade dos dados efetivamente utilizados, que foram tão somente dados de fontes confiáveis e com características precisas, e a correta utilização das variáveis independentes, que apresentaram significância individual reduzida, o que é desejável. Em especial, as limitações na disponibilidade de dados com informações completas para compor a amostra impediu o enquadramento do modelo no Grau III de fundamentação.

Ainda, devido à vasta possibilidade de caracterizar as variáveis explicativas e da diversidade de características que podem ser consideradas na pesquisa, compreendeu-se que este estudo não encerra o assunto, pelo contrário, demonstra apenas uma possibilidade de tratamento para tema. Outros modelos poderão ser elaborados se utilizando desta metodologia, com emprego de mais variáveis ou de diferentes formas de mensuração das mesmas. Um exemplo disso é a variável localização, que é tratada de diversas maneiras pelos avaliadores.

REFERÊNCIAS

- Associação Brasileira de Normas Técnicas. NBR 12721: avaliação de custos unitários de construção para incorporação imobiliária e outras disposições para condomínios edifícios - procedimento. Versão corrigida 2. Rio de Janeiro, 2006. 91p.
- Associação Brasileira de Normas Técnicas. NBR 14653-2: avaliação de bens parte 2: imóveis urbanos. Rio de Janeiro, 2011. 54p.
- Associação Brasileira de Normas Técnicas. NBR 14653-1: avaliação de bens parte 1: procedimentos gerais. Rio de Janeiro, 2001. 10p.
- Dantas, R. A. Engenharia de avaliações: uma introdução à metodologia científica. 3. ed. revisada de acordo com a NBR-14653-2:2011. São Paulo: Pini, 2012. 255p.
- Fiker, J. Manual de avaliações e perícias em imóveis urbanos. 3. ed. São Paulo: Pini, 2008. 157p.