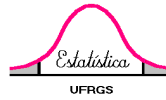




UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL  
INSTITUTO DE MATEMÁTICA  
DEPARTAMENTO DE ESTATÍSTICA



# **Modelo de Regressão Beta para a Análise da Origem dos Problemas de Sistemas Prediais**

Autora: Cecília Brasil Biguelini  
Orientadora: Prof<sup>a</sup> Dr<sup>a</sup>. Márcia Elisa Echeveste

Porto Alegre, 04 de Dezembro de 2009.

Universidade Federal do Rio Grande do Sul  
Instituto de Matemática  
Departamento de Estatística

# Modelo de Regressão Beta para a Análise da Origem dos Problemas de Sistemas Prediais

Autora: Cecília Brasil Biguelini

Monografia apresentada para obtenção  
do grau de Bacharel em Estatística.

Banca Examinadora:

Professora Dr<sup>a</sup> Márcia Elisa Echeveste (orientador)

Professor Dr. Carlos Torres Formoso (co-orientador)

Msc Ângelo Márcio de Oliveira Sant'Anna

Porto Alegre, 04 de Dezembro de 2009.

*Dedico este trabalho a minha avó, Teresinha de Jesus Alcalde Brasil, pelo afeto, amor, conselho e ensinamentos de vida.*

## Agradecimentos

Aos meus pais, Iarema e Vernei, pela dedicação, carinho, ensinamentos, compreensão e por me darem forças para vencer mais uma etapa

Aos meus irmãos, Aloizio e Bruno, pelo carinho e compreensão.

Ao meu namorado Anderson Amaral, pela compreensão, carinho e companheirismo.

À minha orientadora Márcia Echeveste, pela orientação, paciência, estímulo e confiança que apostou em mim

Ao grupo NORIE, Núcleo Orientado para a Inovação da Edificação, em especial Juliana Brito, Letícia Berr e professor Carlos Torres Formoso, pela amizade, contribuições, oportunidades e estímulo a pesquisa,

Ao professor João Riboldi, pelas orientações e conselhos que fizeram com que eu nunca desistisse dos meus objetivos.

A todos que participaram direta e indiretamente desta etapa da minha vida.

## RESUMO

Este artigo apresenta uma aplicação do Modelo de Regressão Beta (MRB) para auxiliar a pesquisa de avaliação de empreendimentos habitacionais, no que diz respeito à associação entre as características físicas destes empreendimentos e as reclamações dos usuários referentes aos problemas relacionados aos sistemas prediais das unidades habitacionais. As reclamações analisadas neste trabalho são 28 empreendimentos habitacionais, localizados no Rio Grande do Sul, as quais foram coletadas a partir de análise documental de ordens de serviço (OS), armazenadas em um sistema de informação de uma empresa administradora participante do Programa de Arrendamento Residencial (PAR). Por meio da análise de aproximadamente 10 mil ordens de serviço foi possível classificar 580 Ordens de Serviço (O.S) e, a partir destas, associar características físicas dos empreendimentos habitacionais que influenciam no número de reclamações de problemas referentes a sistemas prediais em empreendimentos do PAR. As principais contribuições deste estudo são a aplicação de um modelo recentemente discutido na área de estatística, Regressão Beta, para avaliação de empreendimentos habitacionais. Com base nos resultados, é possível verificar quais as características físicas e espaciais dos empreendimentos que impactam na ocorrência de reclamações relacionadas a sistemas prediais realizadas pelos próprios moradores. Tal informação pode contribuir na tomada de decisão e melhoria em projetos futuros.

**Palavras-chaves:** Regressão Beta, PAR, Avaliação de empreendimentos habitacionais.

Este artigo será submetido a “REVISTA Ambiente Construído”

# SUMÁRIO

1. INTRODUÇÃO .....	7
2. SELEÇÃO DO MODELO ESTATÍSTICO .....	8
3. OBJETIVO.....	10
4. PROCEDIMENTOS METODOLÓGICOS .....	11
5. RESULTADOS E DISCUSSÕES .....	12
6. CONSIDERAÇÕES FINAIS .....	20
REFERÊNCIAS BIBLIOGRÁFICAS.....	22

# 1. INTRODUÇÃO

A trajetória da política habitacional tem sido marcada por mudanças na concepção e no modelo de intervenção do poder público, especialmente quanto ao equacionamento do problema da moradia para a população de baixa renda (MINISTÉRIO DAS CIDADES, 2004). Neste contexto, existe a necessidade de promover a melhoria da qualidade e a redução de custos da habitação de interesse social (HIS), proporcionando assim, o adequado atendimento das necessidades dessa população. Isso significa gerar produtos e serviços de qualidade a um custo adequado para todos os envolvidos, principalmente para os moradores, principais usuários da habitação (BRITO 2009).

A busca por melhoria da qualidade dos produtos e serviços e redução de custos dos mesmos ocorre, principalmente, através do reconhecimento das necessidades dos clientes. Na indústria da manufatura é comum a utilização de diversas formas para se captar os requisitos dos clientes, tais como análises das expectativas de clientes em indústrias paralelas, utilização de pessoas disfarçadas em compradores, auditorias externas, pesquisas formais, entre outras. Além dessas, uma tendência emergente dessas indústrias, porém ainda pouco explorada na construção civil, é a utilização de reclamações como fonte de informação para a melhoria da qualidade dos produtos e serviços. Segundo alguns autores, as reclamações constituem um método direto e eficiente para entender os clientes (BARLOW; MOLLER, 1996), possibilitando a melhoria da qualidade dos produtos e serviços.

A utilização de reclamações pode beneficiar o usuário através do atendimento das necessidades e expectativas identificadas, além de auxiliar a empresa na identificação de potenciais pontos de melhoria no produto ou processo (ANG; BUTTLE, 2006; BARLOW; MOLLER, 1996). Desta forma, analisar as reclamações realizadas pelo usuário, além de entender os fatores que os fazem ter a intenção de reclamar, é uma importante forma de aprendizado para a empresa.

Na área de conhecimento relacionada ao marketing, os estudos relativos ao comportamento do consumidor caracterizam o processo de reclamação como dinâmico, ou seja, um processo que depende de uma série de eventos relacionados que podem ser representadas através de uma árvore de decisões. Quando um cliente encontra-se insatisfeito com um produto ou serviço ele pode exercer alguma ação comportamental ou simplesmente afastar-se da empresa. Tal atitude é o primeiro sinal do comportamento



subseqüente do consumidor, a qual pode estar relacionada a ações públicas ou privadas, como ações legais aos órgãos de apoio ao consumidor ou propagandas negativas boca a boca à familiares e amigos (DAY, LANDON, 1977; SINGH, 1988 apud BRITO, 2009).

Da mesma forma, os estudos relativos ao comportamento do consumidor tem se focado na identificação das variáveis que influenciam a intenção de reclamar por parte do consumidor. Essas variáveis, estudadas geralmente no contexto de marketing, podem ter um comportamento diferente no contexto habitacional.

Os modelos estatísticos constituem-se ferramentas úteis para resumir e interpretar dados. Nas pesquisas voltadas à área da habitação de interesse social, as análises realizadas para a avaliação dessas habitações são, na maioria das vezes, de cunho qualitativo, as quais frequentemente não utilizam dados quantitativos para análise. Contudo, muitos bancos de dados são gerados com dados quantitativos de diferentes naturezas como perfil do usuário, dados de respondentes, registro de reclamações e outros dados cuja análise pode contribuir para alavancar a qualidade de projetos futuros. Dessa forma, tem-se a oportunidade de utilizar modelos estatísticos na avaliação de dados provenientes de diferentes etapas do processo de desenvolvimento de empreendimentos habitacionais. Com base neste conhecimento, pode-se proceder ações corretivas e preventivas que permitam melhorar qualidade final de produtos, processos e serviços vinculados a construção de empreendimentos habitacionais. Por meio de pesquisas fundamentadas nos princípios estatísticos, é possível apoiar, com maior robustez, a tomada de decisão por parte dos agentes envolvidos na concepção das formas de provisão habitacional do país.

## 2. SELEÇÃO DO MODELO ESTATÍSTICO

Diante do cenário exposto anteriormente, inicialmente realizou-se uma comparação dos possíveis modelos que podem ser utilizados para detectar quais as características físicas que influenciam na geração de reclamações de empreendimentos habitacionais (problema de pesquisa). Para este caso, a variável resposta de análise é o número de unidades habitacionais que geraram ao menos uma reclamação em relação ao tamanho do empreendimento, isto é, a percentual de unidades reclamantes em relação ao total de unidades habitacionais no empreendimento. Neste trabalho, considerou-se como unidade experimental o empreendimento.

Os modelos de ajuste que podem ser utilizados para esse tipo de variável podem ser encontrados na Figura 1. Como modelos estatísticos possíveis citam-se o Modelo de

Regressão Linear Múltipla (MLR), Modelos Lineares Generalizados, do inglês *General Linear Model (GLM)*, que foram propostos por Nelder & Wedderburn (1972) e o Modelo de Regressão Beta (MRB), que foi proposto por Ferrari & Cribari-Neto (2004).

	Modelo	Distribuição	Função de ligação	Método de estimação	de Intervalo de análise
1	Modelo de Regressão Linear Múltipla	Normal	Identidade	Mínimos Quadrados	$(-\infty, \infty)$
2	Modelo Linear Generalizado	Normal	Log	Máxima Verossimilhança	$(-\infty, \infty)$
3	Modelo de Regressão Beta	Beta	Logit	Máxima Verossimilhança	$[0, 1]$

Figura 1: Resumo dos tipos de modelos de regressão possíveis para o problema de pesquisa

O modelo de Regressão Linear Múltipla é utilizado para descrever o comportamento de uma determinada variável denominadas independentes e na maioria das vezes quantitativa. Segundo Corrar et al. (2007), o modelo baseia-se nos valores de uma ou mais variáveis, denominadas independentes, na maioria quantitativas

O modelo Linear Generalizado é apropriado para investigar o efeito das variáveis independentes sobre uma variável dependente de comportamento não normal. O GLM pertence a uma classe de modelos desenvolvida por Nelder; Wedderburn (1972), a qual se baseia na distribuição de probabilidade com um parâmetro desconhecido ( $\theta$ ), admitindo que a mesma pertença à família exponencial. Este modelo contempla as distribuições Normal, Binomial, Poisson, Gamma, Exponencial e Binomial Negativa e as variáveis independentes do GLM podem ser de natureza quantitativa ou qualitativa. A classe do GLM é definida por um conjunto de variáveis independentes que descreve a estrutura linear do modelo e uma função de ligação entre a média da variável dependente com a estrutura linear do modelo segundo segundo Sant'Anna; Caten (2009).

O modelo de regressão Beta é um tipo de modelo de regressão, cuja variável em questão é restrita a um intervalo unitário ( $0 \leq y \leq 1$ ). Segundo Ferrari; Cribari-Neto (2004) que propuseram o MRB, o modelo permite modelar as relações lineares e não-lineares entre as variáveis explicativas e a variável dependente. Para o problema de pesquisa exposto acima, a variável resposta é proporção e o MRB tem como objetivo modelar este tipo de variável. Segundo Ferrari; Cribari-Neto (2004), o MRB permite uma estimação bastante precisa e segura em relação aos seus coeficientes, independente do tamanho da amostra e da tendência dos dados a valores próximos de zero ou de um.

A família de distribuição Beta é composta de todas as distribuições de probabilidade que apresente uma variável resposta  $Y$  cuja função de densidade depende dos parâmetros  $p$

e  $q$ . Segundo Kieschnick; McCullough (2003) e Ferrari; Cribari-Neto (2004) a distribuição de probabilidade Beta é uma função de distribuição de probabilidade que não pertence à família exponencial, pois sua distribuição não pode ser escrita na forma canônica e apresentar um parâmetro de localização  $\mu$ . A família de distribuições Beta contempla as distribuições Uniforme, Arco-Seno e Dirichlet segundo Sant'Anna; Caten (2009).

Por fim neste trabalho utilizou-se o modelo de regressão Beta, pois é o modelo que melhor se enquadra devido a limitação dos dados, pelo fato da variável em questão ser um percentual. Conforme mencionado anteriormente, ressalta-se que o percentual refere-se a reclamações decorrentes de problemas construtivos em sistemas prediais. O modelo proposto trata-se de um estudo de relacionamento entre variáveis de natureza qualitativa e quantitativa na análise da proporção de um determinado problema.

### 3. OBJETIVO

O objetivo principal deste trabalho é propor a utilização do Modelo de Regressão Beta com finalidade explicar as relações entre as características físicas e espaciais de empreendimentos habitacionais de interesse sociais (EHIS) e o percentual de unidades que realizaram reclamações sobre sistemas prediais<sup>1</sup>.

O Modelo Beta, em geral tem publicações na indústria para área de controle de qualidade com o intuito de controlar a qualidade de atributos de produtos e serviços. Para o estudo de modelos estatísticos, este trabalho apresenta-se como uma oportunidade de aplicação e discussão do modelo na área de construção civil.

Para esta área, o entendimento das características que estão associadas a que tipo de problema (por meio da análise de reclamações dos usuários) é uma informação que evita a recorrência do mesmo em projetos futuros.

---

<sup>1</sup> Os sistemas prediais são sistemas físicos e integrados que tem por finalidade dar suporte às atividades dos usuários, suprindo-os com os insumos prediais necessários e proporcionando os serviços requeridos. Neste trabalho, os sistemas prediais considerados para as análises foram as instalações hidrossanitárias e as instalações elétricas.

## 4. PROCEDIMENTOS METODOLÓGICOS

Os dados utilizados neste trabalho são provenientes da Administradora A que contém dados de 42 empreendimentos do Programa de Arrendamento Residencial (PAR) do Rio Grande do Sul, registradas entre dezembro de 2002 e setembro de 2006 qual atua em 72% dos empreendimentos PAR<sup>2</sup> do Rio Grande do Sul. Os empreendimentos são atribuídos de forma aleatória à administradora, a partir de um processo licitatório para a habilitação das empresas administradoras pela Caixa Econômica Federal.

Neste trabalho, foram analisadas as características: tipologia arquitetônica, tipologia construtiva e a construtora do empreendimento. Ressalta-se que existem outras características que podem estar relacionadas à incidência de reclamações de problemas de sistemas prediais, as quais não foram consideradas neste trabalho por uma limitação dos dados disponíveis.

Este trabalho se caracteriza quanto à natureza por ser uma pesquisa quantitativa, no qual o fenômeno estudado baseou-se na análise de dados quantitativos. O estudo se caracteriza pela averiguação da dependência de uma variável resposta em função de  $k$  variáveis independentes, o que se caracterizaria por relações de dependência univariada múltipla.

Primeiramente, realizaram-se a identificação das reclamações referentes aos problemas relacionados a sistemas prediais. Dentro deste sistema estão incluídos os problemas de não funcionamento de rede elétrica e não funcionamento da rede hidrossanitária (estes dois problemas somam 92,79% dos problemas relacionados a sistemas prediais). Neste trabalho foram analisadas 580 reclamações referentes problemas no sistema predial observados em 28 empreendimentos. Como a variável em questão é um percentual não se pode pressupor que esta siga à distribuição Normal. O modelo de Regressão Beta foi escolhido por possibilitar a previsão de valores dentro do intervalo  $[0,1]$ , o que não acontece no modelo de regressão linear múltipla.

O presente trabalho considerou 28 empreendimentos habitacionais, pois no restante dos casos há uma falha no registro dos dados. Estas, se referem a dados que não foram coletados no ano ine. Ressalta-se que neste trabalho foram consideradas as reclamações realizadas no primeiro ano de ocupação dos empreendimentos. Com base nestas reclamações, obteve-se a razão entre o total de unidades habitacionais com problemas no

---

<sup>2</sup> Até o ano de 2008 o RS possuía 113 empreendimentos PAR. A administradora A administra 81 deles. (dado fornecido pela Administradora A em janeiro de 2009).

sistema predial em relação ao número de unidades habitacionais dos empreendimentos. Com base neste cálculo, obteve-se o percentual de unidades habitacionais com problemas de sistema predial por unidade habitacional.

## 5. RESULTADOS E DISCUSSÕES

Ao analisar o percentual de unidades habitacionais que tem reclamações de sistemas prediais, primeiramente analisou-se a distribuição do percentual de unidades reclamantes referente ao problema citado dos empreendimentos. Na Figura 2 apresenta-se o histograma do percentual dos problemas relacionados a sistemas prediais por empreendimento. Verifica-se através da Figura 2 que os dados não têm uma tendência de serem normalmente distribuídos.

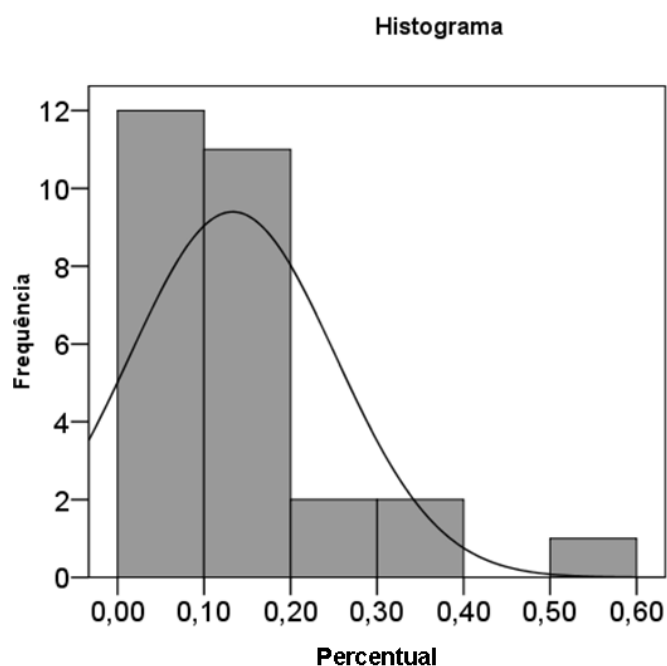


Figura 2: Distribuição do percentual de problemas relacionados a sistemas prediais por empreendimento habitacional

Com base na análise descritiva percebe-se que a média da proporção de unidades habitacionais em relação ao tamanho (número de unidades) do empreendimento é em torno de 14 unidades a cada 100. Através do cálculo da mediana e da moda nota-se que o valor mais provável para a verdadeira proporção está em torno de 11 unidades reclamantes a cada 100. Na Tabela 1 são apresentados os resultados das estatísticas descritivas.

Tabela 1: Estatísticas Descritivas

N	28
Media	0,133
Mediana	0,107
Moda	0,107
Desvio Padrão	0,1189

As variáveis explicativas analisadas no modelo estão classificadas conforme a Tabela 2.

Tabela 2: Características dos empreendimentos

Características dos Empreendimentos	Descrição	Níveis
Tipologia arquitetônica	Classificação do tipo de construção que remete a composição arquitetônica	1 = apartamento 2 = casa/sobrado
Tipologia construtiva	Categoria do programa que define as características da habitação	1 = simplificados 2 = renovação 3 = normal
Construtora	Refere-se a construtora responsável pela obra do empreendimento. No total são 20 construtoras.	C1 a C20

Antes de realizar a análise do modelo de regressão Beta, realizou-se uma análise prévia da variável em questão em relação às variáveis explicativas. Na Figura 3a, as três tipologias construtivas não apresentam diferenças significativas ( $p > 0,05$ ) e na Figura 3b, as tipologias arquitetônicas casa/sobrado e apartamento não se diferem significativamente ( $p > 0,05$ ). No caso das construtoras, há evidências de que há diferença significativa ( $p < 0,01$ ) entre as construtoras a um nível de significância de 5%. A análise estatística utilizada foi a análise de variância (ANOVA) no *software* estatístico SPSS 15.0. A análise de variância (*one way anova*) foi realizada para testar separadamente o quanto cada variável influencia na proporção de unidades reclamantes no empreendimento.

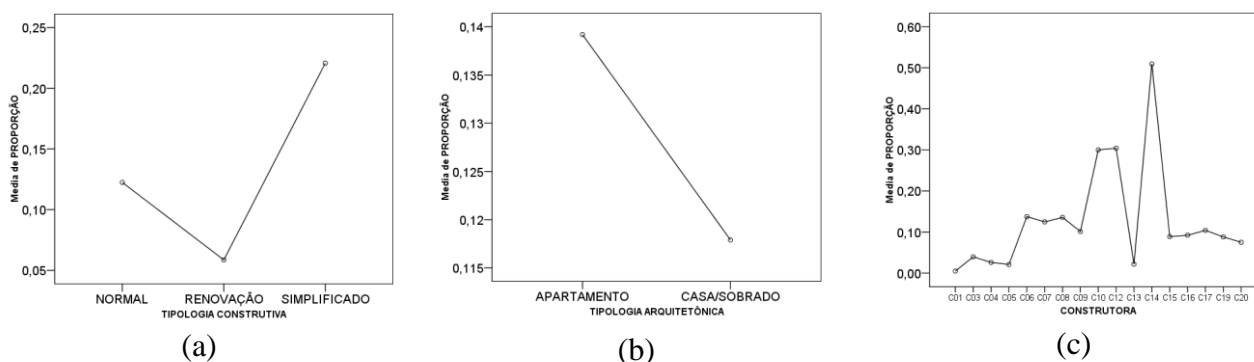


Figura 3: Variável resposta em função das suas características

Para a modelagem do MRB foi necessário fazer uma substituição das variáveis: tipologia construtiva, tipologia arquitetônica e construtora por variáveis *dummy* ( $1 = característica presente$  e  $0 = característica não presente$ ) para a construção do modelo de regressão. As novas variáveis explicativas foram definidas como: tipologia construtiva renovação, tipologia construtiva simplificada e tipologia arquitetônica apartamento e as construtoras como C3, C4, C5, C6, C7, C8, C9, C10, C12, C13, C14, C15, C16, C17, C19 e C20, ou seja, para cada construtora foi gerada uma variável que assumiu valores 0 e 1.

A função de ligação utilizada foi a LOGIT pelo fato desta obter um melhor ajuste do modelo de acordo com os dados. A função de ligação permite modelar a proporção de unidades reclamante do empreendimento habitacional com relação às variáveis explicativas. Esta proporção foi obtida em relação a suas variáveis explicativas através de uma função, representada por  $g(\cdot)$ . Conforme Dobson (2002) a função de ligação  $g(\cdot)$  é responsável pela transformação da média da variável resposta e não de cada observação.

A modelagem de variáveis de natureza qualitativa pressupõe que a estimativa de um dos níveis de cada variável explicativa é nula, enquanto os demais níveis podem assumir valores positivos ou negativos, de acordo com a influência desta sobre a variável resposta. O nível de cada variável explicativa que possui valor nulo torna-se um referencial passível de ser comparado com outros níveis (SANT'ANNA; CATEN, 2005). O nível considerado nulo é sempre o primeiro a ser testado, como, por exemplo, a C1.

No modelo inicial foram utilizadas as três variáveis explicativas citadas anteriormente e os 28 empreendimentos PAR, conforme a Figura 3. Com base na análise de regressão Beta, percebe-se que as variáveis tipologia arquitetônica, tipologia construtiva e construtora foram significativas, porém há um ponto de alavancagem alto neste modelo, conforme a observação referente ao empreendimento PAR20. Ao analisar a Figura 5a, verifica-se este ponto de alavancagem, ou seja, o empreendimento PAR20 foi considerado um ponto discrepante no modelo, e por esse motivo foi retirado da análise. Esta discrepância pode estar relacionada à construtora responsável pela execução deste empreendimento (a construtora do empreendimento foi a C07), a qual possui, em grande parte de seus empreendimentos, um alto número de unidades reclamantes, enquanto que no caso específico do PAR20 tem-se um pequeno número de unidades.

Em relação ao critério de informação de Akaike (AIC), o modelo final foi o modelo que apresentou menor valor de AIC. O critério de AIC é obtido dos estimadores de máxima verossimilhança (EMV), que servem para decidir o modelo que está mais adequado dentre vários modelos de quantidades diferentes de coeficientes, ou seja, auxilia

na seleção do modelo melhor ajustado de acordo com o número reduzido de coeficientes. Neste caso, o modelo final foi o que melhor se ajustou.

Ao analisar os critérios de adequação dos modelos, percebe-se que o coeficiente de determinação ajustado do primeiro modelo foi de 92,19%. Este índice demonstra que as variáveis consideradas conseguem explicar em 92,19% a proporção de unidades habitacionais que apresentam reclamações de sistemas prediais, enquanto que no modelo final as variáveis explicativas conseguem explicar 96,57% da variável resposta. Tabela 3 abaixo apresenta a estimativa dos coeficientes do modelo de regressão Beta, juntamente com o erro padrão, a significância de cada variável e os critérios de adequação. Para realizar as análises de Regressão Beta foi utilizado o software R2. 8.2.

Tabela 3: Estimativa dos coeficientes, erros padrões e critérios de adequação do modelo

Parâmetro	Estimativa	Erro Padrão	Significância	Exp(Estimativa)
Intercepto	-4,9375	0,6905	<0,001	
<i>Tipologia Construtiva</i>				
Normal	-	-	-	-
Renovação	-1,1680	0,2557	<0,001	0,3110
Simplificado	0,7577	0,4946	0,1260	2,1333
<i>Construtora</i>				
C1	-	-	-	-
C3	1,3590	0,7967	0,0880	3,8922
C4	1,4221	0,8139	0,0806	4,1458
C5	1,1887	0,8122	0,1430	3,2828
C6	3,1153	0,7171	<0,001	22,5402
C7	3,1482	0,6995	<0,001	23,2941
C8	2,3427	0,8704	0,0071	10,4093
C9	3,2045	0,7122	<0,001	24,6432
C10	3,3369	0,8610	<0,001	28,1318
C12	4,1208	0,7059	<0,001	61,6085
C13	1,2552	0,8283	0,1300	3,5085
C14	4,2167	0,8591	<0,001	67,8093
C15	2,6006	0,7328	<0,001	13,4718
C16	3,1004	0,7135	<0,001	22,2068
C17	2,0474	0,8757	0,0194	7,7477
C19	2,6491	0,7507	<0,001	14,1413
C20	3,6210	0,7781	<0,001	37,3749
<i>Tipologia Arquitetônica</i>				
APARTAMENTO	-	-	-	-
CASA/SOBRADO	-0,0270	0,1803	0,8810	0,9734
Null Deviance	666,8455(26)			
Residual Deviance	26,6734(7)			
R <sup>2</sup> ajustado	0,9657			
AIC	-97,9585			



O modelo final de previsão do MRB é descrito na equação (1).

$$g(\mu) = - 4,9375 - 1,1680\text{renovação} + 0,7577\text{simplificado} + 1,3590C3 + 1,4221C4 + 1,1887C5 + 3,1153C6 + 3,1482C7 + 2,3427C8 + 3,2045C9 + 3,3369C10 + 4,1208C12 + 1,2552C13 + 4,2167C14 + 2,6006C15 + 3,1004C16 + 2,0474C17 + 2,6491C19 + 3,6210C20 - 0,0270\text{CASA/SOBRADO} \quad (1)$$

Ao avaliar as estimativas do modelo final, observa-se que para o modelo de regressão Beta, a estimativa de risco para tipologia construtiva renovação resultou na equação (2).

$$\exp(\text{renovação}) = \exp(-1,1680) = 0,3109 \quad (2)$$

Em outras palavras, quando é escolhida a variável tipologia construtiva renovação, a chance de que haja uma reclamação relacionada a sistema predial diminui em 31% em relação a tipologia construtiva normal. Em relação a tipologia construtiva simplificado, esta apresenta 2,13 vezes mais chances de ter uma reclamação por problema sobre sistemas prediais em relação a tipologia construtiva normal.

De acordo com os resultados, o modelo estima que a variável tipologia arquitetônica CASA/SOBRADO, possui 97,34% menos chance de ter reclamações de problemas construtivos relacionados ao sistema predial se comparada a tipologia de APARTAMENTOS.

No caso da construtora do empreendimento, a construtora C1 é a construtora que será comparada com as outras, ou seja, as construtoras que apresentam uma chance maior da ocorrência de reclamações em seus empreendimentos são C14 (67,81), C12(61,61), C20(37,38) e C10(28,61) e as construtoras que tem menos chances de que haja reclamações são C5(3,28), C13(3,51), C3(3,89) e C4(4,15) em relação a construtora C1. A construtora C14 tem 67,81 vezes mais chance de produzir reclamações do que a construtora C1.

A Figura 4 apresenta o gráfico de resíduo *deviance* do modelo inicial e o do modelo final. Os dois modelos apresentam uma aleatoriedade nos resíduos *deviance*, ambos não apresentam tendência nas distribuições ou pontos aparentemente discrepantes. Com base nisso, verifica-se que a função de ligação está adequada. Segundo Atkinson e Riani (2000), o resíduo *deviance* é obtido com duas vezes a diferença entre o máximo do logaritmo da verossimilhança do modelo nulo e do modelo da pesquisa. Testa-se o valor do resíduo *deviance* com o ponto crítico  $\chi^2(u-k)(\alpha)$  da distribuição Qui-quadrado com significância igual a  $\alpha$  (usualmente  $\alpha < 0,05$ ),  $n$  número de observações e  $k$  coeficientes do modelo. Caso o resíduo for menor ao ponto crítico encontrado, pode-se considerar que há evidências que

o modelo analisado esteja bem ajustado aos dados. Caso contrário, descarta-se o modelo, pois o mesmo pode ser considerado inadequado.

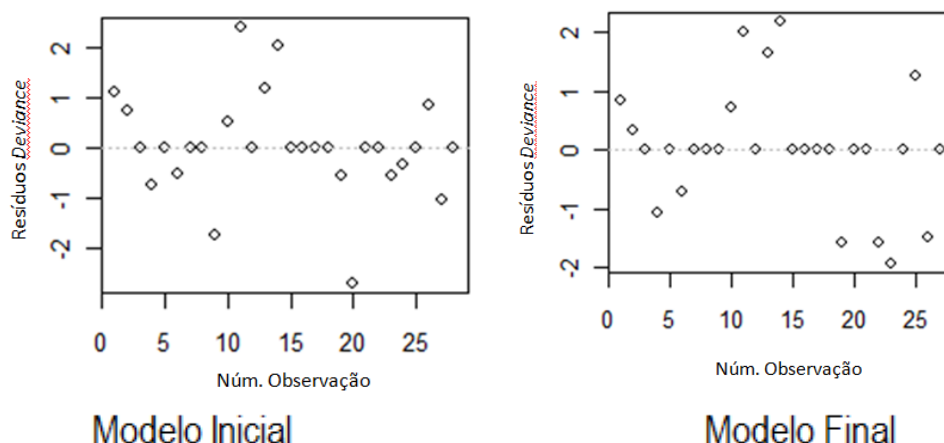


Figura 4: Gráfico de Resíduos *Deviance* para o modelo inicial e o modelo final

Para detectar a influência de cada observação nas estimativas dos coeficientes do modelo de regressão, utilizou-se a distância de Cook apresentada no modelo inicial e no modelo final. Para ambos os modelos, as observações apresentam valores muito altos para a distância de Cook ( $4 \times 10^{28}$  no modelo final), uma vez que o valor máximo desejável é de 1. As altas distâncias possivelmente estejam associadas ao pequeno tamanho da amostra. Dessa forma, o modelo gerado nesta pesquisa pode ser melhorado com uma amostra mais significativa, pois neste caso o modelo apresentado não se pode detectar a influência de cada observação nas estimativas dos coeficientes com precisão.

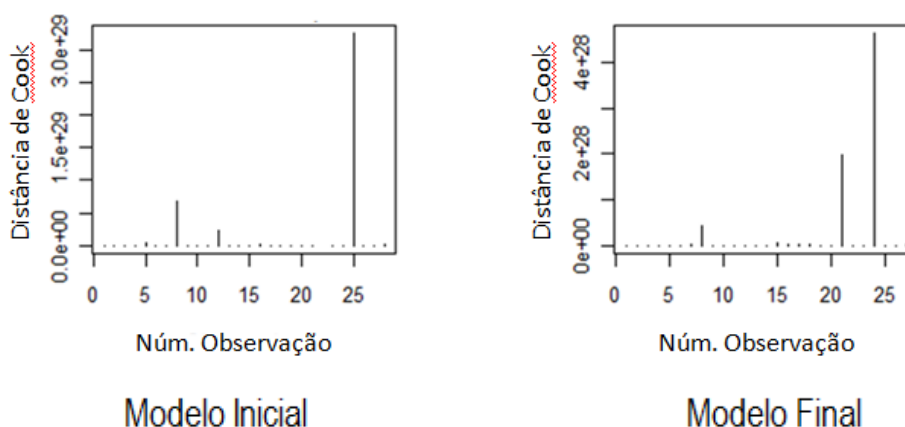


Figura 5: Gráfico de Distância de Cook para o modelo inicial e o modelo final

Para medir a influência das observações em modelos de regressão, avaliando a importância individual de cada observação no próprio valor ajustado, utilizou-se a alavancagem generalizada. Essa ferramenta demonstra graficamente a ocorrência de pontos de alavanca no modelo.

Com relação aos pontos de alavancagem generalizada, verificando através do modelo inicial, com a retirada do empreendimento PAR 20, percebe-se que no modelo final não há pontos altos de alavancagem. O modelo final foi selecionado, pois além de não possuir mais pontos de alavancagem, o critério de informação de AIC foi reduzido, tornando o modelo final mais adequado.

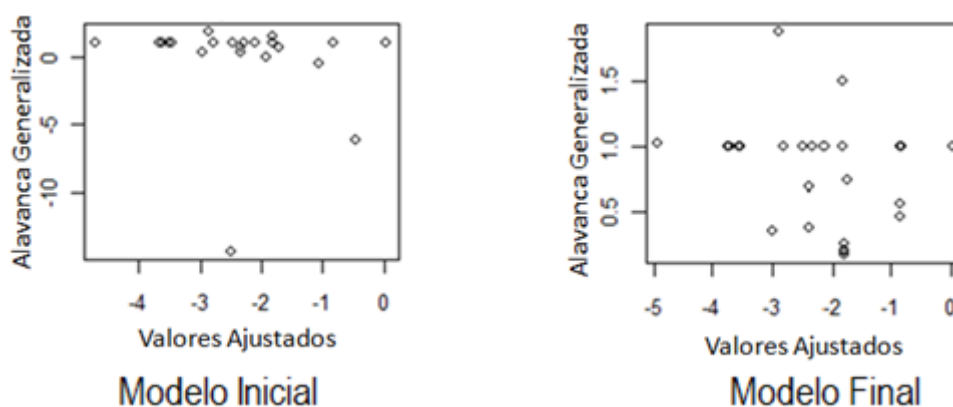


Figura 6: Gráfico da Alavanca Generalizada para o modelo inicial e o modelo final

Após o resultado da distancia de Cook e considerando o baixo tamanho de amostra, alternativamente, as 17 construtoras foram reagrupadas em 4 categorias de acordo com , Para verificar se o fato de existirem muitos níveis de construtoras podem estar influenciando nos critérios de adequação do Modelo de Regressão Beta, agrupou-se as 17 construtoras em 4 níveis conforme apresentado na Tabela 4. O critério utilizado para agrupá-los foi o percentual de problemas dos empreendimentos.

<i>Classificação</i>	<i>Nível</i>
Até 0,05	Muito Bom
0,05 a 0,10	Bom
0,10 a 0,15	Regular
Acima de 0,15	Ruim

Tabela 4: Níveis para o agrupamento de construtoras

Analisando as estimativas dos parâmetros na Tabela 5, percebe-se que as variáveis construtora, tipologia construtiva simplificado ( $p < 0,04$ ) influenciam no percentual de unidades reclamantes dos empreendimentos relacionadas a sistemas prediais.

Tabela 5: Estimativa dos coeficientes, erros padrões e critérios de adequação do modelo com variável construtora agrupada

Parâmetro	Estimativa	Erro Padrão	Significância
Intercepto	-2,0508	0,3456	<0,001
<i>Tipologia Construtiva</i>			
Normal	-	-	-
Renovação	-0,4450	0,4943	0,3680
Simplificado	0,5969	0,2921	0,0410
<i>Construtora</i>			
Muito Bom	-1,2444	0,5234	0,0174
Bom	-	-	-
Regular	-0,0635	0,3883	0,8700
Ruim	0,7130	0,3205	0,0261
<i>Tipologia Arquitetônica</i>			
APARTAMENTO	-0,6806	0,3912	0,0819
CASA/SOBRADO	-	-	-
Null Deviance	65,7771(27)		
Residual Deviance	24,1908(20)		
R <sup>2</sup> ajustado	0,6351		
AIC	-66,8980		

Verificando o coeficiente de determinação com as construtoras agrupadas, verifica-se que as variáveis explicativas conseguem explicar 63,5% do percentual de unidades reclamantes nos empreendimentos, ou seja, outras variáveis explicativas podem estar influenciando na variável em questão. Em relação ao critério de adequação de AIC, o modelo com mais níveis de construtoras se mostrou melhor adequado, ou seja, o modelo anterior possuía menor AIC.

Na Figura 4, apresentam-se os gráficos de adequabilidade do modelo com as construtoras agrupadas: gráfico de resíduos, gráfico da alavanca generalizada e distancia de cook. No gráfico dos resíduos *deviance* percebe-se que os resíduos estão aleatoriamente distribuídos e no gráfico da alavanca generalizada não há evidências de nenhum ponto discrepante no modelo. Em relação a distância de Cook, com o agrupamento dos níveis da variável construtora, nota-se que a distancia das observações encontra-se dentro do limite aceitável (abaixo de 1).

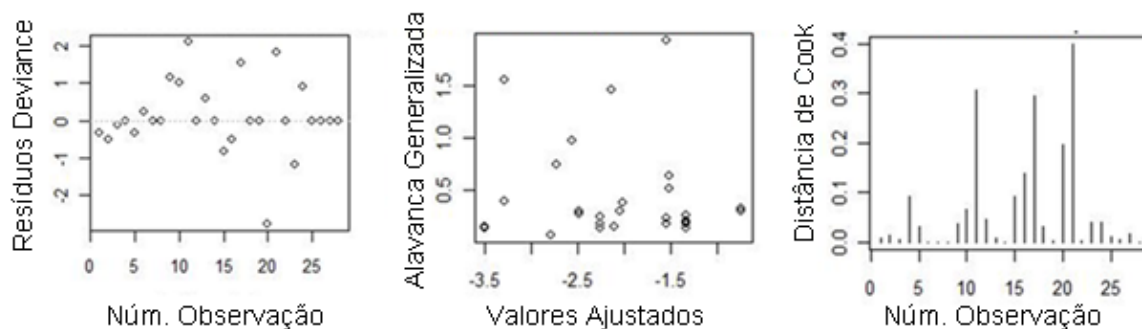


Figura 4: Gráficos de Adequação do Modelo

O modelo com variáveis agrupadas mostrou-se adequado. Ressalta-se que a informação por construtora é possível de ser levantada com base no modelo original (com as 17 construtoras), contudo o tamanho da amostra deveria ser superior.

Em outras palavras, o modelo original deve ser replicado para um número maior de observações, pois a precisão da análise seria melhor e poderia servir como tomada de decisão para empreendimentos futuros e evitar recorrência das reclamações.

O tipo de construção é outra variável apontada como significativa no modelo, o que pode estar associado com a tipologia arquitetônica dos empreendimentos. As casas e sobrados são, na maioria das vezes, PAR simplificado, enquanto que os apartamentos são PAR normal. Essa relação pode influenciar na ocorrência da maior média de reclamações de sistemas prediais nos edifícios de apartamentos, pois a complexidade das instalações são maiores, possibilitando a maior ocorrência dos mesmos.

## 6. CONSIDERAÇÕES FINAIS

A realização do estudo da etapa de uso de EHIS tornou possível identificar características importantes para a tomada de decisão nas etapas do desenvolvimento de projetos habitacionais. A presente pesquisa apresentou a aplicação de um modelo de regressão recentemente consolidado na área de estatística como alternativa para verificar a influência de alguns fatores em relação a problemas construtivos por empreendimento. Contudo, o estudo original realizado não apresentou qualidades preditivas, possivelmente em razão do tamanho da amostra. Uma sugestão para trabalhos posteriores é utilizar um processo construtivo com amostra maior, gerando uma precisão melhor que possibilite a tomada de decisão em futuros empreendimentos habitacional de interesse social. Com o agrupamento das construtoras reduzindo o número de categorias, o modelo atestou qualidades preditivas.

Constatou-se neste trabalho que as reclamações de sistemas prediais podem estar relacionadas com a construtora e com o tipo de construção do empreendimento. As informações obtidas através do modelo de regressão, dado que se tenha um modelo adequado, com amostra considerável, podem auxiliar os tomadores de decisão no ato da contratação de determinada empresa construtora. O modelo é capaz de estimar a porcentagem média de unidades habitacionais nos empreendimentos que reclamam em função da construtora analisada.

Este trabalho foi desenvolvido com apoio de órgãos financiadores de estudos em habitação de interesse social. Ressalta-se que o uso de modelos matemáticos vem crescendo como alternativa a criação de indicadores capazes de avaliar os empreendimentos habitacionais. O trabalho evidenciou que é possível identificar as relações entre características de projetos e as reclamações resultantes, porém outras associações são possíveis dependendo das variáveis disponíveis para a realização das análises. Tal procedimento é fundamental para a proposição de melhoria na qualidade da habitação, bem como no desenvolvimento de ações capazes de minimizar a recorrência de problemas advindos da má qualidade, incorporando essas análises na concepção do projeto de edificações. Neste cenário, a inclusão do uso do modelo de regressão Beta para identificar as características dos empreendimentos que estão influenciando nos problemas construtivos mostrou-se uma técnica adequada.

Por fim, ressalta-se que outros tipos de reclamações podem ser testados através do Modelo de Regressão Beta, possibilitando uma contribuição mais ampla para o entendimento da variável resposta que podem influenciar cada tipo de reclamação. No presente trabalho, três características foram consideradas no modelo (construtora, tipo de construção e tipologia arquitetônica). Dessa forma, sugere-se que outras análises sejam realizadas considerando outras características.

## REFERÊNCIAS BIBLIOGRÁFICAS

- ATKINSON, A. C.; RIANI, M. **Robust Diagnostic Regression Analysis**, Nova York: Springer, 2000.
- BARLOW, J.; MOLLER, C. **Reclamação de Cliente? Não tem Melhor Presente**: usando o feedback do cliente como uma ferramenta estratégica. São Paulo: Futura, 1996.
- BIGUELINI, C.B.; et al. Desenvolvimento de Modelos Lineares Gerais (GLM) para o estudo das reclamações de usuários de empreendimentos habitacionais em função das características físicas dos empreendimentos. In: 54° REUNIÃO ANUAL DA REGIÃO BRASILEIRA DA SOCIEDADE INTERNACIONAL DE BIOMETRICA, 13° SIMPÓSIO DE ESTATÍSTICA APLICADA À EXPERIMENTAÇÃO AGRONÔMICA, 54, 13., 2009, São Carlos. **Anais...** São Carlos: UFSCAR/EMBRAPA, 2009.
- BRITO, J. N. S. **Retroalimentação do Processo de Desenvolvimento de Empreendimentos de Habitação de Interesse Social a Partir de Reclamações de Usuários**: estudo no Programa de Arrendamento Residencial. 2009. Dissertação (Mestrado em Engenharia Civil) – Programa de Pós-Graduação em Engenharia Civil, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2009.
- CAIXA ECONÔMICA FEDERAL. **Termo de Referência Programa de Arrendamento Residencial**.– Brasília: GEPAD/DIDUP, 2003.
- CORRAR, L. J; PAULO, E; DIAS FILHO, J. M. **Análise multivariada**: para cursos de administração, ciências contábeis e economia. São Paulo: Atlas, 2007.
- CORDEIRO, G.M. **Modelos Lineares Generalizados**, In: SIMPÓSIO NACIONAL DE PROBABILIDADE E ESTATÍSTICA, 7., 1986, Campinas, SP. **Anais...** São Paulo: ABE, 1986. 286 p.
- CORDEIRO G. M.; DEMÉTRIO C. G. B. **Modelos Lineares Generalizados**. Minicurso para o 12° SEAGRO e 52° Reunião Anual da RBRAS, USFM, Santa Maria, RS, 2007.
- FERRARI, S. L. P.; CRIBARI-NETO, F. Beta Regression for Modeling Rates and Proportions. **Journal of Applied Statistics**, v. 31, p. 799–816, 2004.
- KIESCHNICK, R.; MCCULLOUGH, B. D. Regression Analysis of Variates Observed on (0,1): percentages, proportions and fractions, **Statistical Modelling**, v. 3, p. 193–213, 2003.
- MCCULLAGH, P.; NELDER, J. A. **Generalized Linear Models**. 2. ed. Londres: Chapman & Hall, 1989.
- MONTGOMERY, D. C.; PECK, E. A.; VINING, G. G. **Introduction to Linear Regression Analysis**. 4. ed. Nova York: John Wiley & Sons, 2006.
- NELDER, J. A.; WEDDERBURN, R. W. M. Generalized Linear Models. **Journal of the Royal Statistical Society A**, v. 135, n. 3, p. 370-84, 1972.
- OLIVEIRA, M. S. **Um Modelo de Regressão Beta**: teoria e aplicação. Dissertação (Mestrado em Matemática). Instituto de Matemática e Estatística, Universidade de São Paulo, São Paulo, SP, 2004.
- PAULA, G. A. **Modelos de Regressão com Apoio Computacional**. São Paulo: IME/USP, 2004.
- RAO, C. R.; WU, Y. Linear Model Selection by Cross-Validation. **Journal Statistical Planning and Inference**, v. 128, p. 231–240, 2005.
- SANT'ANNA, A. M. O. **Método de Orientação à Modelagem de Dados Mensurados em Proporção**. Dissertação (Mestrado em Engenharia de Produção) – Escola de Engenharia, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2006.
- Sant'Anna, A.M.O. & Caten, C.S. (2009) Modelagem da fração de não-conformes em processos industriais, **Pesquisa Operacional**, no prelo.