Cleberton Correia Santos
(Organizador)

Estudos Interdisciplinares nas Ciências e da Terra e Engenharias

Atena
Editora
Ano 2019

Cleberton Correia Santos
(Organizador)

Estudos Interdisciplinares nas Ciências Exatas e da Terra e Engenharias

Atena Editora
2019

Editora Executiva: Profª Drª Antonella Carvalho de Oliveira
Diagramação: Natália Sandrini
Edição de Arte: Lorena Prestes
Revisão: Os Autores

Dados Internacionais de Catalogação na Publicação (CIP)
(eDOC BRASIL, Belo Horizonte/MG)

E82 Estudos interdisciplinares nas ciências exatas e da terra e engenharias 1 [recurso eletrônico / Organizador Cleberton Correia Santos. – Ponta Grossa, PR: Atena Editora, 2019. – (Estudos Interdisciplinares nas Ciências Exatas e da Terra e Engenharias; v. 1)

Formato: PDF
Requisitos de sistema: Adobe Acrobar Reader
Modo de acesso: World Wide Web
Inclui bibliografia
ISBN 978-85-7247-621-8
DOI 10.22533/at.ed.218191109

1. Ciências exatas e da Terra. 2. Engenharias. 3. Tecnologia. I.Santos, Cleberton Correia. II. Série.

CDD 016.5

Elaborado por Maurício Amormino Júnior – CRB6/2422

Atena Editora
Ponta Grossa – Paraná - Brasil
www.atenaeditora.com.br
contato@atenaeditora.com.br

Atena Editora
Ano 2019

APRESENTAÇÃO

O livro “**Estudos Interdisciplinares nas Ciências Exatas e da Terra e Engenharias**” de publicação da Atena Editora apresenta em seu primeiro volume 35 capítulos relacionados temáticas de área multidisciplinar associadas à Educação, Agronomia, Arquitetura, Matemática, Geografia, Ciências, Física, Química, Sistemas de Informação e Engenharias.

No âmbito geral, diversas áreas de atuação no mercado necessitam ser elucidadas e articuladas de modo a ampliar sua aplicabilidade aos setores econômicos e sociais por meio de inovações tecnológicas. Neste volume encontram-se estudos com temáticas variadas, dentre elas: estratégias regionais de inovação, aprendizagem significativa, caracterização fitoquímica de plantas medicinais, gestão de riscos, acessibilidade, análises sensoriais e termodinâmicas, redes neurais e computacionais, entre outras, visando agregar informações e conhecimentos para a sociedade.

Os agradecimentos do Organizador e da Atena Editora aos estimados autores que empenharam-se em desenvolver os trabalhos de qualidade e consistência, visando potencializar o progresso da ciência, tecnologia e informação a fim de estabelecer estratégias e técnicas para as dificuldades dos diversos cenários mundiais.

Espera-se com esse livro incentivar alunos de redes do ensino básico, graduação e pós-graduação, bem como pesquisadores de instituições de ensino, pesquisa e extensão ao desenvolvimento estudos de casos e inovações científicas, contribuindo então na aprendizagem significativa e desenvolvimento socioeconômico rumo à sustentabilidade e avanços tecnológicos.

Cleberton Correia Santos

SUMÁRIO

CAPÍTULO 19

AVALIAÇÃO PELA MODA, MÉDIA OU MEDIANA?

Luiz Fernando Palin Droubi
Universidade Federal de Santa Catarina
Florianópolis - Santa Catarina
Norberto Hochheim
Universidade Federal de Santa Catarina
Florianópolis - Santa Catarina
Willian Zonato
Universidade Federal de Santa Catarina
Florianópolis - Santa Catarina

RESUMO: Este artigo trata do problema da retransformação de variáveis aplicada no caso de utilização da transformação logarítmica da variável dependente. A compreensão deste problema é útil não apenas na engenharia de avaliações, área de trabalho dos autores, mas também na Economia e nas mais diversas áreas em que a inferência estatística seja utilizada e onde se aplique a transformação logarítimica à variavel dependente, em busca de conseguir um melhor modelo de regressão. Especificamente na Engenharia de Avaliações, contudo, o debate sobre qual medida de tendência central deve ser utilizada no processo de retransformação do modelo de regressão tomou grandes proporções, e provoca polêmica e conflitos inclusive em perícias judiciais envolvendo arbitramento de valores de imóveis, seja para desapropriações, seja para resolução de outros conflitos jurídicos de qualquer natureza em que se necessite da avaliação de um imóvel, como na execução de uma garantia, numa liquidação forçada ou mesmo por questões inventariais.
PALAVRAS-CHAVE: retransformação, regressão, tendência central.

ABSTRACT: This article deals with the problem of the retransformation of variables applied in case of using the logarithmic transformation of the dependent variable. The understanding of this problem is useful not only in the Appraisals Engineering, our working area, but also in Economics and in the most diverse areas where statistical inference is used and where logarithmic transformation is applied to the dependent variable, in order to obtain a better regression model. Specifically in the Appraisals Engineering, however, the debate about what measure of central tendency should be used in the retransformation problem has taken on great proportions, and causes controversy and conflicts especially in judicial investigations involving arbitration of property values, either for expropriations or for resolution of other legal conflicts of any nature in which the assessment of a property is required, such as the execution of a guarantee, a forced liquidation or even for inventory matters.
KEYWORDS: retransformation, regression, central tendency.

1 I INTRODUÇÃO

Existe na área da avaliação de imóveis uma discussão frequente e a nosso ver indesejável a respeito da estimativa de tendência central adotada para a predição de valores quando da utilização de modelos lineares log-normais, isto é, modelos em que a variável resposta aparece transformada pela função logaritmo natural.

Como será visto oportunamente, quando um modelo linear log-normal for homocedástico (σ = cte) e estiver razoavelmente bem ajustado, com um baixo erro-padrão, a adoção de qualquer estimativa de tendência central, moda, média ou mediana, resultará em valores praticamente equivalentes, com variação dentro da precisão da área de avaliações imobiliárias. No entanto, na presença de grande dispersão dos dados, o valor do erro-padrão da regressão linear pode se tornar relativamente alto e a diferença entre as avaliações por uma ou outra medida de tendência central pode tornar-se relevante, levando a uma situação altamente indesejável: um imóvel poderá ser avaliado por dois avaliadores independentes com uma diferença significativa entre os valores encontrados. Tendo em vista que a NBR14.653-02 (2011) se omite a este respeito, as duas avaliações serão válidas, porém com valores altamente discrepantes.

Pretende-se com este artigo dar a este problema uma abordagem formal, com o intuito de sugerir uma padronização das avaliações, sem no entanto especificar qual medida de tendência central é a correta, haja vista que todas elas são matematicamente válidas, apresentando prós e contras, nenhuma delas podendo ser considerada melhor que a outra.

2 I DESENVOLVIMENTO E FUNDAMENTAÇÃO

> Major Point 1: When we talk about the relationship of one variable to one or more others, we are referring to the regression function, which expresses the mean of the first variable as a function of the others. The key word here is *mean*! (MATLOFF, 2009, p. 386, grifo do autor)

2.1 Valor Esperado

Segundo BENNETT (2006), a **esperança matemática** ou **valor esperado** de uma variável aleatória é a soma do produto de cada probabilidade de saída da experiência pelo seu respectivo valor. Isto é, representa o valor médio 'esperado' de uma experiência se ela for repetida muitas vezes. Matematicamente, a Esperança de uma variável aleatória X false é representada pelo símbolo E(X).

Segundo Matloff (2009, p. 43), o valor esperado tem um papel central em probabilidade e estatística. A definição mais ampla de valor esperado de uma variável aleatória X, válida tanto para variáveis discretas como contínuas, é:

$$\lim_{n\to\infty} = \frac{X_1 + \ldots + X_n}{n}$$

2.1.1 Cômputo do Valor Esperado de uma variável aleatória discreta

Segundo Matloff (2009, p. 44), o valor esperado de uma variável aleatória *X* que assume valores definidos no conjunto *A* é:

$$E(X) = \sum cP(x = c)$$

onde P(X= c) representa a função probabilidade da variável aleatória X assumir o valor *c*.

2.1.2 Cômputo do Valor Esperado de uma variável aleatória contínua

O Valor Esperado de uma variável aleatória contínua *W* pode ser escrito da seguinte forma (MATLOFF, 2009, p. 128):

$$E(W) = \int_{-\infty}^{\infty} t f_W(t)\,dt$$

onde f_w (t) é a função densidade de probabilidade de *t*, para todo *t* onde a função f_w (t) esteja definida.

2.1.3 Propriedades do Valor Esperado

Seja *a* um escalar e *U* uma variável aleatória (MATLOFF, 2017, p. 47):

$$E(aU) = aE(U)$$

Sejam *a* e *b* dois escalares e *U* e *V* duas variáveis aleatórias, não necessariamente independentes, então (MATLOFF, 2017, p. 47):

$$E(aU + bV) = aE(U) + bE(V)$$

Finalmente, sejam *U* e *V* duas variáveis aleatórias ***independentes***:

$$E(UV) = E(U)E(V)$$

Porém, se U e V não forem independentes, esta propriedade falha (covariância).

2.2 Lei da expectativa total

Segundo Matloff (2009, p. 339), a lei da expectativa total pode ser expressa como abaixo:

$$E(Y) = E[E(Y \mid W)]$$

2.3 Lei da Variância total

Outra relação importante é expressa pela lei da Variância Total que, de acordo com Matloff (2009, p. 345) estabelece que:

$$VAR(Y) = E[VAR(Y \mid W)] + VAR[E(Y \mid W)]$$

2.4 Desigualdade de Jensen

Segundo Matloff (2017), se φ(x) é uma função convexa, então a desigualdade de Jensen se exprime na seguinte desigualdade:

$$\varphi(E(X)) \leq E(\varphi(X))$$

Como pode-se demonstrar, a função e^x é uma função convexa, pois possui derivada segunda sempre maior que zero ($f''(e^x) = e^x > 0$)).

2.5 Erro médio quadrático (MSE)

Seja π o valor de uma estimativa. Então o seu erro médio quadrático (MSE) é dado por:

$$MSE = \int (y-\pi) f(y)dy = E(y^2) - 2\pi E(y) + \pi^2$$

Para encontrar o valor mínimo do erro médio quadrático (MSE) quando π varia, faz-se:

$$d\frac{(E(Y^2)-2\pi E(Y)+\pi^2)}{d\pi}=0\leftrightarrow\pi=E(Y)$$

Ou seja, a estimativa pelo valor esperado é a estimativa que minimiza e erro médio quadrático.

2.6 Valor Esperado condicional

O valor esperado de uma variável aleatória *Y* estatisticamente relacionada com outra outra variável aleatória *X* é (WASSERMAN, 2010, p. 77):

$$E(Y\mid X=t)=\int tf_{Y|X}(Y\mid X=t)dt$$

2.7 Estimadores

> Earlier, we often referred to certain estimators as being "natural." For example, if we are estimating a population mean, an obvious choice of estimator would be the sample mean. But in many applications, it is less clear what a "natural" estimate for a population quantity of interest would be. We will present general methods for estima- tion in this section. We will also discuss advanced methods of inference (MATLOFF, 2009, p. 303).

A definição de um *estimador* Ɵ para um parâmetro ou uma variável é uma função $\hat{\theta}(X)$, que mapeia o espaço amostral para um conjunto de estimativas amostrais, em que *X* é uma variável aleatória dos dados observados. É usual denotar uma estimativa em um determinado ponto x є X por $\hat{\theta}(X=x)$ ou, mais simplesmente, $\hat{\theta}(x)$.

2.8 Propriedades de um estimador

Nesta secção adota-se como notação que é $\hat{\theta}$ um estimador da variável aleatória Ɵ.

2.8.1 Erro

$$e(x)=\hat{\theta}(x)-\theta$$

2.8.2 Desvio

$$d(x)=\hat{\theta}(x)-E[\hat{\theta}(x)]$$

onde $E[\hat{\theta}(x)]$ é o Valor Esperado do estimador.

2.8.3 Variância

A variância de um estimador $\hat{\theta}$ é (MATLOFF, 2009, p. 52):

$$VAR(\hat{\theta}) = E[(\hat{\theta} - E(\hat{\theta}))^2]$$

2.8.4 Coeficiente de Variação

O coeficiente de variação de um estimador é uma medida adimensional que compara o desvio-padrão de uma variável ou estimador θ à sua média, conforme abaixo (MATLOFF, 2009, p. 56):

$$CV = \frac{VAR(\hat{\theta})}{E(\hat{\theta})}$$

2.8.5 Viés

O viés de um estimador $\hat{\theta}$ é (MATLOFF, 2009, p. 317):

$$B(\hat{\theta}) = E(\hat{\theta}) - \theta$$

O viés coincide com o valor esperado do erro, pois $E(\hat{\theta}) - \theta = E(\hat{\theta} - \theta)$. Numa regressão linear:

$$B(\hat{\mu}(x_0)) = E[\hat{\mu}(x_0)] - \mu(x_0)$$

2.8.6 Erro médio quadrático

Segundo Shen e Zhu (2008, p. 553), o erro médio quadrático é uma medida comum da qualidade de um estimador na literatura estatística.

$$MSE = E[(\hat{\theta} - \theta)^2]$$

Numa regressão linear, o erro médio quadrático pode ser descrito por:

$$MSE[\hat{\mu}(x_0)] = E[(\hat{\mu}(x_0) - \mu(x_0))^2] = VAR[\hat{\mu}(x_0)] + B^2[\hat{\mu}(x_0)]$$

2.8.7 Consistência

A consistência é a propriedade que um estimador tem de se aproximar assintoticamente do valor “real” da variável, a medida que aumenta o número de observações. Matematicamente:

$$\lim_{n \to \infty} \hat{\theta} = \theta$$

2.8.8 Melhor estimador linear não-viesado ou BLUE

Em estatística, é comum o uso da sigla BLUE (*Best Linear Unbiased Estimator*) para indicar o melhor estimador linear não-viesado.

2.9 Trade-off entre viés e variância

Um dos problemas conhecidos dos modelos de regressão linear ou outros modelos estatísticos em geral é o sobreajustamento (do inglês *overfitting*). Resumidamente, *overfitting* é o ato de ajustar um modelo tão bem ajustado aos dados amostrais, que este se torna incapaz de fazer boas previsões para outros dados que não os do próprio modelo. Segundo Matloff (2017, p. 24), um modelo sobreajustado é um modelo tão elaborado que “capta o ruído ao invés do sinal”.

De acordo com Matloff (2017, pp. 24–26), por outro lado, um modelo com menor número de variáveis explicativas estará enviesando os seus resultados (no sentido de enviesamento sistêmico, inerente à amostragem, não proposital), e o acréscimo de uma variável independente a este modelo estará assim reduzindo o seu viés.

Por outro lado, de acordo com Matloff (2017, p. 25), quanto maior for o número de variáveis do modelo – mantido o mesmo número de dados amostrais –, maior será a variabilidade coletiva dos regressores e, assim, maior será a variância dos coeficientes estimados.

Desta maneira, em modelos mais simples, a redução do viés do mesmo através da adição de um novo regressor compensa o aumento na variabilidade conjunta do modelo, até que este número de regressores atinja um número ótimo, quando a diminuição adicional do viés gerada pela adição de um regressor torna-se tão pequena que não compensa a variabilidade dos coeficientes estimados. Um modelo com variáveis explicativas maior do que este número ótimo estará, portanto, sobreajustado.

Ou seja, existe um ***tradeoff*** entre *viés* e *variância*: para qualquer estimador estatístico (MATLOFF, 2017, p. 25), não se pode reduzir o seu viés sem aumentar a sua variância e vice-versa. Tem-se que conviver sempre com algum viés e tem que se aceitar alguma variância.

Matematicamente, isto decorre do desenvolvimento da expressão do Erro Médio

Quadrático (MSE) (MATLOFF, 2017, p. 49):

$$MSE(\hat{\theta}) = E[(\hat{\theta} - \theta)^2] = E[\hat{\theta} - E(\hat{\theta}) + E(\hat{\theta}) - \theta]^2$$

Desenvolvendo a expressão acima, chega-se:

$$MSE(\hat{\theta}) = E[(\hat{\theta} - E(\hat{\theta}))^2] + E[(E(\hat{\theta}) - \theta)^2] + E[2(\hat{\theta} - E(\hat{\theta}))(E(\hat{\theta}) - \theta)]$$

como:

- o termo $E[(\hat{\theta} - E(\hat{\theta}))^2]$ é igual à variância do estimador ($VAR(\hat{\theta})$);
- o termo $E[(E(\hat{\theta}) - \theta)^2]$ é o quadrado do viés do estimador ($B^2(\hat{\theta})$);
- e, finalmente, o termo $E[2(\hat{\theta} - E(\hat{\theta}))(E(\hat{\theta}) - \theta)]$ é nulo, haja vista que $E[\hat{\theta} - E(\hat{\theta})] = 0$.

Matematicamente, portanto, temos que:

$$MSE(\hat{\theta}) = VAR(\hat{\theta}) + B^2(\hat{\theta})$$

2.10 Regressão Linear

2.10.1 Definição precisa

Sejam *Y* e *X* duas variáveis e $m_{Y;X}(t)$ uma função tal que:

$$m_{Y;X}(t) = E(Y \mid X = t)$$

Chama-se $m_{Y;X}$ de **função de regressão de *Y* dado *X*** (MATLOFF, 2009, p. 386, grifo do autor). Em geral, $m_{Y;X}(t)$ é a **média** de *Y* para todas as unidades da população para as quais $X = t$ (MATLOFF, 2009, p. 386, grifo nosso).

> The word “regression” is an allusion to the famous comment of Sir Francis Galton in the late 1800s regarding “regression toward the mean.” This referred to the fact that tall parents tend to have children who are less tall closer to the mean – with a similar statement for short parents. The predictor variable here might be, say, the father’s height *F*, with the response variable being, say, the son’s height *S*. Galton was saying that $E(S|F) < F$.

Segundo Matloff (2009, p. 386, grifo do autor), ainda, a função $m_{Y;X}(t)$ é uma função da **população**, ou seja, apenas **estima-se** uma equação de regressão ($\hat{m}_{Y;X}(t)$) à partir de uma amostra da população.

> The function $m_{Y;X}(t)$ is a population entity, so we must estimate it from our sample data. To do this, we have a choice of either assuming that $m_{Y;X}(t)$ takes on some parametric form, or making no such assumption. If we opt for a parametric approach, the most common model is linear [...] (MATLOFF, 2009, p. 389).

Segundo Matloff (2009, pp. 394–397), as proposições acima sobre a função $m_{Y;X}$ podem ser generalizadas para outras quantidades de regressores em *X* e seus termos de interação, tal que:

$$m_{Y;X}(t) = \beta_0 + \beta_1 t_1 + \beta_2 t_2 + \beta_3 t_1 t_2 + \beta_4 t_1^2$$

Notando que o termo **regressão linear** não necessariamente significa que o gráfico da função de regressão seja uma linha reta ou um plano, mas que se refere a função de regressão ser linear em relação aos seus parâmetros (β_i)

2.10.2 Estimação em modelos de regressão paramétricos

Segundo Matloff (2009, p. 389), é possível demonstrar que o mínimo valor da quantidade1 $E[Y - g(X))^2]$ é obtido, entre todas as outras funções, para $g(X) = m_{Y;X}(X)$. Porém, "se pretendemos minimizar o erro médio absoluto de predição, E(IY - g(X)I), a melhor função seria a mediana *g*(*Y*) = mediana(*Y*I*X*)." (MATLOFF, 2009, p. 389).

Matloff (2009) aqui está implicitamente se referindo a um outro tipo de regressão, chamada de regressão quantílica, mais especificamente, à regressão à mediana, ou seja, ao quantil de 50%.

2.10.3 A equação de regressão linear

Como será visto nesta secção, a equação de regressão linear μ(t) é uma *função da população*, que geralmente não nos está acessível, pois se tem acesso a não mais do que uma parte (amostra) desta população em estudo. O que usualmente se faz, então, é *estimar* uma equação de regressão $\hat{\mu}(t)$ para que se possa prever os valores reais da variável em análise.

Tem que se levar em conta que a equação de regressão linear não é uma equação determinística, mas probabilística. No dia-a-dia da prática de engenharia de avaliações, assim como em outras áreas, no entanto, a equação de regressão é usualmente escrita simplificadamente, sem o termo de erro ϵ, ou seja, a equação de regressão é escrita como uma equação determinís- tica, da forma *Y* = α + *X*β ou, exemplificando em termos de variáveis de avaliação de imóveis, VU = α + *A*β, onde VU representa o valor unitário dos imóveis e *A* a sua área.

No entanto, estas equações são uma simplificação da equação de regressão. Na verdade, a equação de regressão μ(t) é uma função da *população* e pode ser escrita formalmente como abaixo (MATLOFF, 2017, p. 66):

$$\mu(t) = \beta_0 + \beta_1 t_1 + \ldots + \beta_p t_p$$

Como o termo de erro da equação, ou seja, o erro que seria cometido ao prever *Y* se a equação de regressão da população fosse efetivamente conhecida, é (MATLOFF, 2017, p. 67):

$$\varepsilon = Y - \mu(t)$$

Então pode-se escrever a equação de regressão de outra maneira, como abaixo (MA- TLOFF, 2017, p. 67):

$$Y = \beta_0 + \beta_1 t_1 + \ldots + \beta_p t_p + \varepsilon$$

Onde ϵ é uma variável aleatória supostamente tal que E(ε) = 0 e VAR(ε) = σ^2, ou simplesmente ε ~ N(0,σ^2).

Num modelo onde não há a adoção de qualquer transformação para a variável dependente, verificada a hipótese da normalidade, esta equação de regressão é também a equação de estimação da variável *Y*, ou seja, para uma equação de regressão sem transformação de variáveis, pode-se escrever:

$$E(Y \mid X) = E(\alpha + X\beta) + E(\varepsilon) = \alpha + X\beta$$

Haja vista que o valor esperado para o termo de erro ε é igual a zero.

No entanto, quando a variável dependente *Y* é transformada, este mesmo termo de erro, desprezado na equação de regressão acima, é de suma importância para o computo do valor esperado da variável original, como será visto neste artigo, pois ele determina a equação de estimação da variável original. Por exemplo, no caso que aqui nos interessa, que é o da transformação logarítmica da variável dependente, tem-se:

$$\ln(Y) = \alpha + X\beta + \varepsilon \Leftrightarrow$$
$$Y = \exp(\alpha + X\beta).\exp(\varepsilon) \Leftrightarrow$$
$$E(Y \mid X) = E[\exp(\alpha + X\beta)]E[\exp(\varepsilon) \mid X] <$$
$$E(Y \mid X) = \exp(\alpha + X\beta).E[\exp(\varepsilon) \mid X]$$

O fundamental a se perceber aqui é que, quando há transformação da variável dependente, para voltarmos à variável original, tem que se levar em conta o termo de erro, haja vista que uma propriedade do valor esperado é a de que $E(f(X)) \neq f(E(X))$, como será visto a seguir.

Mais precisamente, para funções convexas, pela ***desigualdade de Jensen***, $f(E(X)) \leq E(f(X))$ .Isto implica que o valor esperado da exponencial do termo de erro que precisa-se estimar é maior do que a exponencial do valor esperado do erro, ou seja, $E(expe(\varepsilon) \mid X) \geq exp(E(\varepsilon \mid X))$, ou $E(exp(\varepsilon) \mid X) \geq 1$.

Desta maneira, não é correto imaginar que a equação de estimação da variável original esteja acessível pela simples retransformação (pela função exponencial) dos coeficientes da equação de regressão logaritmizada.

Ou seja, a consideração de que os erros aleatórios e com distribuição normal na equação de regressão logaritmizada possam ser diretamente retransformados por um fator de erro multiplicativo igual a 1 é equivocada, já que isto viola a desigualdade de Jensen. O fator de erro multiplicativo, pela desigualdade de Jensen, é maior do que 1.

Desta maneira, não seria correto afirmar que, ao utilizar a avaliação pela média, se esteja violando os princípios clássicos da regressão linear, já que na verdade o que ocorre é justamente o contrário: ao utilizar a média, estamos reafirmando a equação de regressão, que inclui o termo de erro.

GIANNAKOS; LEÃO (1996) faz uma crítica à utilização da avaliação pela moda da distribuição lognormal, crítica esta muito bem elaborada e da qual não se discorda no todo. Concorda-se que a moda não é o valor mais provável, contudo, pelo motivo que **o valor mais provável é o Valor Esperado** da variável, ou seja, o seu valor médio, como será visto oportunamente. E encontra-se mesmo em GIANNAKOS; LEÃO (1996), que “a média aritmética é o ‘valor esperado’ da variável”. Porém, o mesmo trabalho faz também uma defesa da utilização da estimativa pela mediana desta distribuição, o que não estaria de acordo com o que prevê a teoria.

Mesmo a avaliação pela média da variável lognormal não é exata, haja vista que inexiste um estimador exato na inferência estatística. O que existe são estimadores com maior ou menor ***viés***, e maior ou menor ***variância***, ou ainda, existem estimadores com maior ou menos ***Erro Médio Quadrático***. Na verdade, o que poderia ser afirmado é que, ao avaliar pela média, o avaliador estaria se aproximando melhor da equação de regressão do que ao avaliar pela moda ou pela mediana, haja vista que faz parte da equação de regressão o termo de erro multiplicativo, de valor sabidamente maior do que 1.

2.10.4 O problema da retransformação das variáveis

De acordo com Shen e Zhu (2008, p. 552), modelos lineares lognormais tem muitas aplicações e muitas vezes é de interesse prever a variável resposta ou estimar a média da variável resposta na escala original para um novo conjunto de covariantes.

Segundo Shen e Zhu (2008, p. 552), se $Z = (Z_1, ..., Z_n)^T$ é o vetor variável resposta de distribuição lognormal e $x_i = (1, x_{i1}, ..., x_{ip})^T$ é o vetor dos covariantes para a observação i, um modelo linear lognormal assume a seguinte forma:

$$Y = \ln(Z) = X\beta + \varepsilon$$

onde $X= (X_1, ..., X_n)^T$, $\beta = (\beta_0, \beta_1, ..., \beta_p)^T$, e $\varepsilon = (\varepsilon_1, ..., \varepsilon_n)^T$ com $\varepsilon_1 \sim N(0,\sigma^2)$ i.i.d.(*identically independently distributed*) (SHEN; ZHU, 2008, pp. 552–553).

Em muitos casos, para um novo conjunto de covariantes x_0 , pode-se estar interessado em prever a variável resposta em sua escala original:

$$Z_0 = e^{x_0^T\beta+\varepsilon_0}$$

ou estimar a média condicional da variável resposta:

$$\mu(x_0) = E[Z_0 \mid x_0] = e^{x_0^T\beta+\frac{1}{2}\sigma^2}$$

De acordo com Shen e Zhu(2008, p. 553), se β e σ^2 são ambos conhecidos, então é fácil demonstrar que o melhor estimador de Z_0 é de fato $\mu(x_0)$. Contudo, na prática, ambos β e σ^2 são desconhecidos e precisam ser estimados para a obtenção de $\mu(x_0)$.

Segundo Shen e Zhu (2008, p. 552), existem na literatura diversos estimadores baseados em diversos métodos inferenciais, como **ML** (*Maximum Likelihood Estimator*), **REML** (*Restricted ML Estimator*), **UMVU** (*Uniformly Minimum Variance Unbiased Estimator*), além de um estimador **REML** com viés corrigido.

Na prática, estes estimadores pertencem a uma classe de estimadores definida na ex- pressão abaixo:

$$\{\hat{\mu}_c(x_0) : \hat{\mu}_c(x_0) = \exp(x_0^T \hat{\beta} + xRSS/2), c = \frac{1}{n-a}, a < n\}$$

Shen e Zhu(2008) então propõem dois novos estimadores baseados na minimização do erro médio quadrático assintótico (*MM*) e do viés assintótico (*MB*).

De maneira que a diferença entre os estimadores supra-citados pode ser resumida ao parâmetro α:

$$\begin{aligned} a_{ML} &= 0 \\ a_{REML} &= p+1 \\ a_{MM} &= p-1-3nv_0-3RSS/(2m) \\ a_{MB} &= p+1-nv_0-RSS/(2m) \end{aligned}$$

2.10.5 Estimadores não-paramétricos

De acordo com Duan (1983, p. 606), o Valor Esperado E de uma variável resposta *Y* que tenha sido transformada em valores η durante a regressão linear por uma função g (Y) **não-linear** não é igual ao valor da simples retransformação da variável transformada pela sua função inversa h(η) = g^{-1} (Y). Em outros termos (DUAN, 1983, p. 606):

$$E(Y_0) = E(h(x_0\beta + \varepsilon)) \neq h(x_0\beta)$$

Reparar que o termo de erro faz parte da composição do valor esperado da variável de regressão. Em uma regressão linear clássica, sem transformação, E (ε) = 0, então $E(Y_0) = E(x_0\beta)$.

Numa regressão linear logaritmizada, ou seja, uma regressão linear com o logaritmo da variável dependente (h (η) = g^{-1} (η) = exp(η)), para efetuar apropriadamente a retransformação das estimativas de volta a sua escala original, precisa-se ter em conta a desigualdade mencionada na seção 2.4.

Segundo (MANNING; MULLAHY, 1999), quando se ajusta o logaritmo natural de uma variável *Y* contra outra variável *X* através da seguinte equação de regressão:

$$\ln(Y) = \beta_0 + \beta_1 X + \varepsilon$$

Se o erro ε é normalmente distribuído, com média zero e desvio padrão σ^2 , então (DUAN, 1983, p. 606; MANNING; MULLAHY, 1999, p. 6):

$$E[Y \mid X] = e^{\beta_0+\beta_1 X} . E[e^{\varepsilon}] \neq e^{\beta_0+\beta_1 X}$$

Embora o valor esperado dos resíduos ε seja igual a zero, ele está submetido a uma transformação não linear, de maneira que não se pode afirmar que $E[e^{\varepsilon}]$ = 1 (como foi visto na seção 2.4, E[exp(X)] > exp(E[X]). Desta maneira, o estimador abaixo, chamado em (SHEN; ZHU, 2008, p. 554) de *naive back-transform estimator*, ou simplesmente **BT** não é consistente e é enviesado, tendo viés multiplicativo de valor assintótico igual a $e^{-\sigma^2/2}$:

$$BT = E[Y \mid X] = e^{\beta_0+\beta_1 X}$$

Segundo (SHEN; ZHU, 2008, p. 554), ainda, o valor de $e^{\sigma^2/2}$ é sempre menor do que 1(SHEN; ZHU, 2008, p. 554).

> As a result, the BT estimator underestimates μ(*x*0), and the bias is large when σ^2 is large. In our study, it appears that the BT estimator performs much worse than the other estimators[…]Actually, the BT estimator is more suitable for estimating the median of Z_0, which is $\exp(x^T_0\beta)$in this case.

Porém se o termo de erro ϵ é normalmente distribuído $N(0,\sigma^2)$, então um estimador não-enviesado para o valor esperado E(Y), de acordo com DUAN (1983), assume a forma vista na equação abaixo(DUAN, 1983, p. 606; MANNING; MULLAHY, 1999, p. 2 e 6):

$$E(Y) = e^{\beta_0+\beta_1 X} . e^{\frac{1}{2}\sigma^2}$$

Cabe salientar, segundo (MANNING; MULLAHY, 1999, p. 6), que se o termo de erro não for i.i.d (independente e identicamente distribuído), mas for homoscedástico, então:

De qualquer maneira, o valor esperado de *Y* é proporcional à exponencial da previsão na escala log.

DUAN (1983) apresenta então um estimador não-paramétrico (*smearing estimate*), independente da função de transformação h(η) e da distribuição dos erros F(ε), tal que:

$$E(\hat{Y}_0) = \int h(x_0\hat{\beta} + \varepsilon) = \frac{1}{n}\sum_{i=1}^{n} h(x_0\hat{\beta} + \hat{\varepsilon}_i)$$

2.11 Considerações sobre os valores de σ*

Segundo Limpert (LIMPERT et al., 2001, p. 346), distribuições lognormais de diversas ciências tem, em geral, valores de s* (o desvio padrão da amostra, na escala original) variando de 1,1 a 33 (na escala logarítmica, entre 0,095 e 3,497), sendo que o mais comum é que estes valores estejam entre 1,4 e 3 ($0,336 \leq s \leq 1,099$ na escala logarítmica).

2.12 Modelos Heteroscedásticos

Modelos heteroscedásticos não são raros, especialmente no caso de variáveis envolvendo valores em moeda, sendo muito comum em modelos econométricos. Em sua essência, são heteroscedásticos aqueles modelos lineares cujo termo de erro não pode ser considerado totalmente independente, ou seja, existe alguma função (linear ou não), tal que $E(e^{\varepsilon}) = f(X)$, de modo que:

$$\ln(E[Y \mid X]) = X\beta + \ln(f(X))$$

É desnecessário dizer que, para estes modelos, o estimador para a média é diferente de $E(Y) = e^{\beta_0+\beta_1 X}.e^{\frac{1}{2}\sigma^2}$, haja vista que, neste caso, σ^2 não é mais um escalar, mas uma função.

Existem diversas maneiras de se contornar este problema. Por exemplo, através da eliminação do viés através da utilização de uma função que modele a variância $\sigma^2(X)$, ou através do estimador sanduíche.

Cabe ainda salientar que, para os modelos heteroscedásticos, não apenas os erros estão comprometidos, mas também os intervalos de confiança.

2.13 Validação Cruzada

Em inferência estatística é fundamental que os erros sejam avaliados não apenas sobre o conjunto amostral dos dados do modelo, mas também que o modelo encontrado efetue boas previsões para novos conjuntos de dados, afinal, na engenharia de avaliações, o intuito final é estimar o valor de um *novo* imóvel, baseado num conjunto amostra de dados semelhantes ao avaliando.

Validação Cruzada ou *cross-validation* é uma técnica estatística que pode ser utilizada de diversas maneiras e consiste em dividir um conjunto de dados em duas ou mais partições distintas, chamados de partição ou partições de treino (*training set*) e partição de teste(*test set*), utilizadas para o ajuste do modelo e para a previsão da variável dependente, respectivamente. Os dados previstos na partição de teste são então comparados aos valores observados.

Neste artigo será utilizada a validação-cruzada utilizando o procedimento chamado de *delete-one procedure*, ou *leave-one-out*, em que se retira apenas um dado do conjunto de dados, ajusta-se um modelo e então utiliza-se este modelo para prever o valor da variável dependente para o dado retirado (SHEN; ZHU, 2008, p. 564).

Para cada observação então calcula-se o seu erro quadrático ($_{Yi}$ - $\hat{Y}i)^2$, utilizado para o cálculo da estatística **RMSPE** (erro de previsão médio quadrático, ou *root mean squared prediction error*) conforme expressão a seguir (SHEN; ZHU, 2008, p. 564):

$$RMSPE=\left(\frac{1}{n}\sum_{i=1}^{n}(Y_i-\hat{Y}_i)^2\right)^{1/2}$$

3 I ESTUDO DE CASO

Com o fim de averiguar qual estimador melhor se adequa ao procedimento de retransfor- mação de variáveis, aplicar-se-á um comparativo entre os estimadores média, moda e mediana, através do uso da estatística RMSPE.

3.1 Procedimento com dados randômicos

Foram realizadas diversas simulações com dados randomicamente gerados através do software **R** versão 3.5.1.

Conforme mencionado na seção 2.5, os valores de s* já encontrados nas aplicações práticas encontram-se entre 1,1 e 33 (0,095 e 3,497), sendo mais comum que estejam entre 1,4 e 3 (0,336 ≤ *s* ≤ 1,099). As simulações, portanto, serão feitas para valores de s* dentro destes limites.

Nas figuras 1 e 2 podem ser vistos graficamente os modelos para alguns valores de σ* , na escala logarítimica e na escala original.

Para estes modelos, cujas principais estatísticas estão ilustradas na tabela 1, pode ser visto na figura 3, como variam as estimativas realizadas com cada um deles através de 3 estimadores: moda, média e mediana. Observa-se que, como os modelos são semelhantes, *i.e.* apresentam praticamente os mesmos coeficientes, diferindo apenas no erro-padrão, o valor da mediana permanece praticamente inalterado, enquanto moda e média, que para baixos valores de σ* praticamente coincidem com a estimativa da mediana, se afastam progressivamente desta quando os valores de σ* aumentam.

	Dependent variable:				
	log(y) (1)	log(y1) (2)	log(y2) (3)	log(y3) (4)	log(y4) (5)
x	0,125***	0,125***	0,126***	0,126***	0,130***
	(0,0003)	(0,001)	(0,001)	(0,002)	(0,003)
Constant	0,001	0,005	−0,083	−0,031	−0,304
	(0,017)	(0,044)	(0,075)	(0,136)	(0,205)
Observations	200	200	200	200	200
R^2	0,999	0,994	0,982	0,942	0,885
Adjusted R^2	0,999	0,994	0,982	0,942	0,885
Residual Std. Error (df = 198)	0,095	0,242	0,417	0,757	1,138

Tabela 1: Comparação dos diversos modelos gerados, com diferentes erro-padrão

Note: *p<0,1; **p<0,05; ***p<0,01

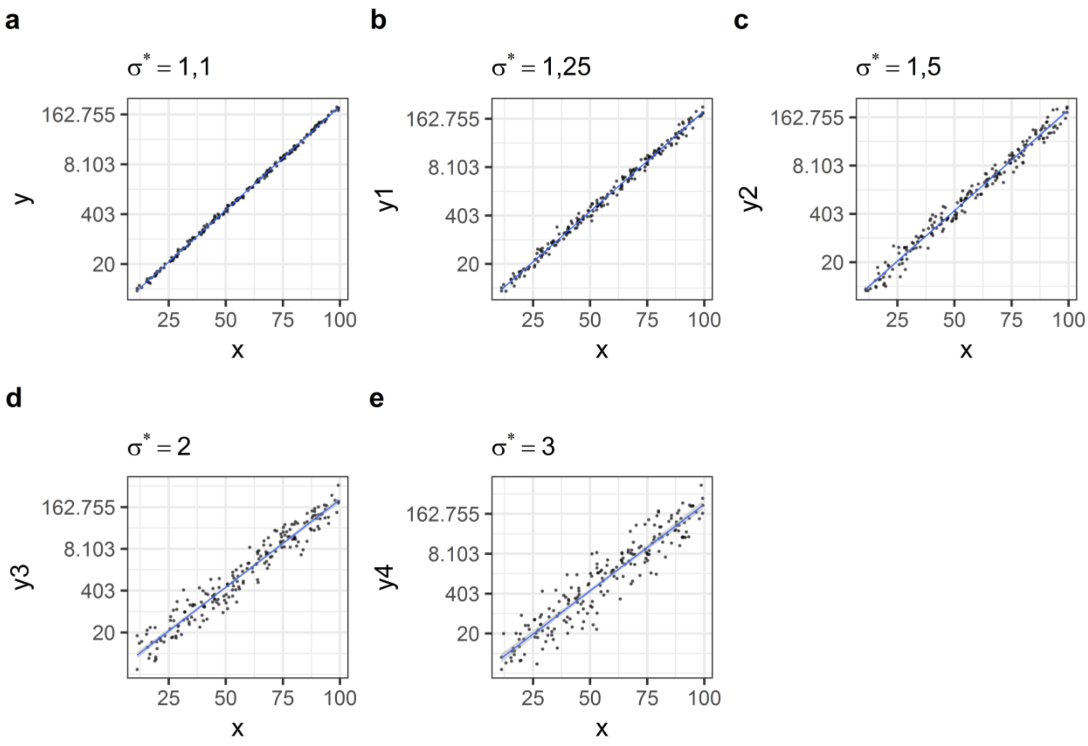

Figura 1: Diversas regressões similares, com diferentes valores de erro-padrão (escala logarítmica.

Fonte: Autores.

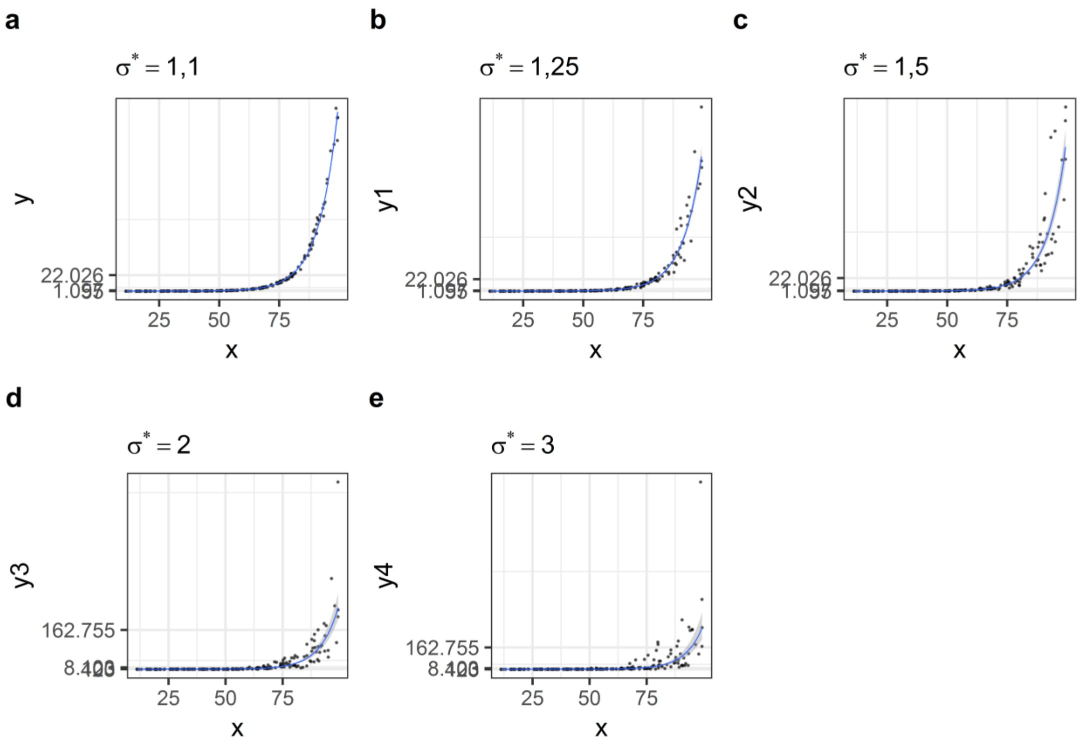

Figura 2: Diversas regressões similares, com diferentes valores de erro-padrão (escala original).

Fonte: Autores.

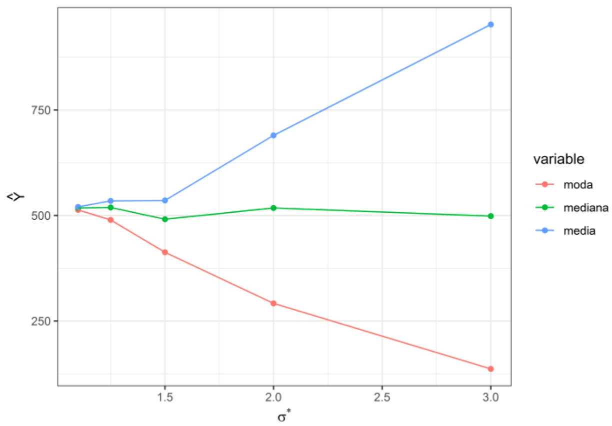

Figura 3: Impacto do erro-padrão no cálculo da estimativas segundo as diversas medidas de tendência central.

Fonte: Autores.

Foram gerados, então, randomicamente, 200 dados uniformes variando de 10 a 100 para variável independente e 200 dados lognormais para a variável dependente, estatisticamente cor- relacionados com a variável independente, tal que o erro padrão ajustado da equação de regressão ln(Y) ~ X varie de 1,1 a 33, em passos de 0,1. Para cada valor de erro padrão ajustado, foram gerados 500 modelos de regressão linear, utilizando-se 70% (140) dos dados escolhidos randomicamente (partição de treinamento), efetuando-se as estimativas sobre os 30% (60) dos dados restantes (partição de testes). Para os dados da partição de testes, então, foi calculado o RMSPE para os diversos estimadores (média, moda e mediana).

Devido à aleatoriedade da escolha das partições de testes e treinamento, o menor RMSPE pode estar tanto na moda, como na média ou na mediana, dependendo dos dados escolhidos.

Nas figuras 4 e 5, podem ser vistos o número de vezes em que cada uma das estimativas obteve o menor valor de RMSE entre elas, quando varia o erro-padrão da regressão.

Percebe-se claramente na figura 4 que, para baixos valores de erro-padrão, a média predomina como melhor estimativa. À partir de um valor de erro-padrão aproximadamente igual a 5, a mediana torna-se a estimativa com menor RMSE.

Na figura 5, pode-se ver os resultados das simulações, porém apenas para a faixa de valores dita mais comum ($1,4 \leq \sigma^* \leq 3$), onde percebe-se que sempre a média tem um melhor comportamento.

Nesta faixa, pelas simulações, a estimação pela média obteve maior eficiência do que a estimação pela mediana ou pela moda, ou seja, os valores de RMSE para as estimativas pela média são menores do que os estimados pela moda ou mediana em aproximadamente 50 a 60% dos casos (250/300 em 500).

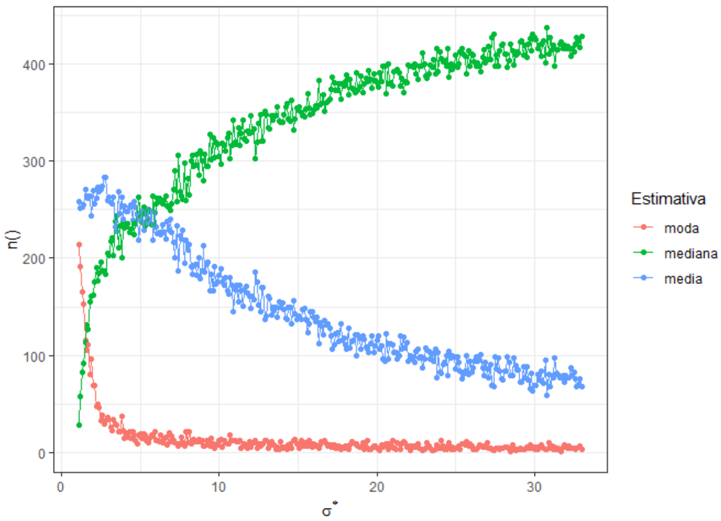

Figura 4: Resultados das simulações.

Fonte: Autores.

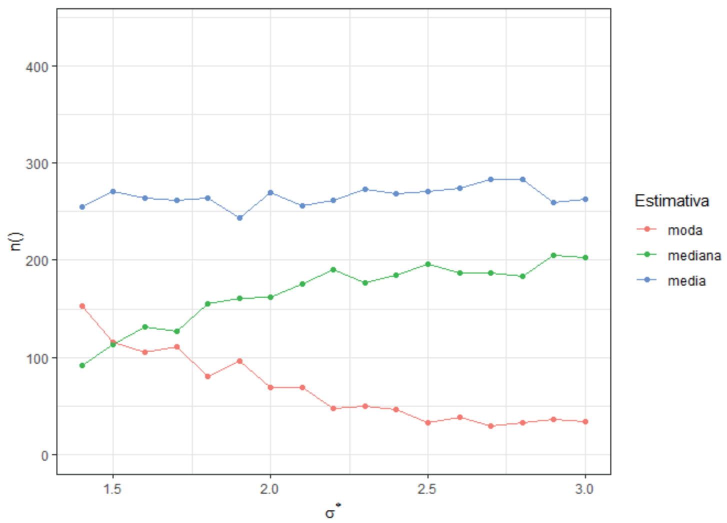

Figura 5: Resultados das simulações para valores mais normais de erro-padrão.

Fonte: Autores.

3.2 Regressão com dados reais de mercado

3.2.1 Dados

Neste estudo será comparada a precisão de diversos modelos estatísticos sobre dados de mercado reais disponíveis em Hochheim (2015, pp. 21–22) e com dados gerados aleatoriamente. A distribuição da variável dependente (valor), pode ser vista na figura 6.

Pode-se mostrar que os dados de Hochheim (2015, pp. 21–22) utilizados no estudo de caso deste artigo, de acordo com a estimação **MLE** (*Maximum Likelihood Estimator*), possuem média $\overline{x}^* = 789.611{,}2$ e desvio-padrão s* = 1,851, calculadas conforme Limpert (2001, p. 345) e o modelo encontrado na mesma referência possui erro-padrão $\hat{\sigma} = 0{,}136$. Para valores de $\hat{\sigma}$ tão baixos como este, as estimativas efetuadas com a média, moda ou mediana são praticamente idênticas, com variação de mais ou menos 1 ou 2% entre as estimativas. Porém, para valores apenas um pouco mais altos de $\hat{\sigma}$, verifica-se que a tendência é que a diferença entre as estimativas realizadas por estes diferentes estimadores se tornem relevantes, o que será mostrado oportunamente.

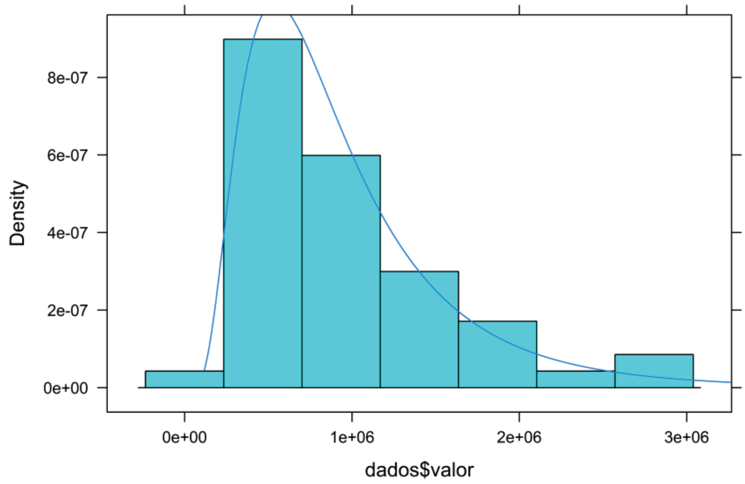

Figura 6: Histograma da variável valor e sua distribuição teórica (lognormal).

Fonte: Autores.

3.2.2 Cálculo do RMSPE

Para o cálculo do RMSPE foi utilizado como referência o modelo proposto por Hochheim(2015, p. 29), ou seja, foram utilizadas as mesmas transformações de variáveis utilizadas no modelo proposto. Os valores dos $\hat{\beta}_i$ são calculados a cada passo.

Os valores encontrados para o erro de predição médio quadrático para cada estimador foram: **R$203.939,11** para a média, **R$204.006,84** para a mediana e **R$205.537,36** para a moda.

Como esperado, o RMSPE foi menor para a média, e maior para a moda. O que comprova a teoria, já que o *naive estimator* é viesado, com viés conhecido de $-\sigma^2/2$, logo a moda possui viés de $-1{,}5\sigma^2$.

3.2.3 Impacto do erro-padrão

Na tabela 2 são mostrados, analogamente ao que foi feito no exemplo anterior com dados randômicos, os valores calculados das estimativas pela moda, média e mediana para o bem avaliando (ver HOCHHEIM (2015), p.22) pelo modelo de Hochheim (2015, p. 29), com o erro- padrão do modelo (0,136) e para outros valores de erro-padrão.

Pela análise dos valores da tabela, percebe-se que, para diversos modelos com iguais coeficientes $\hat{\beta}_i$, mas com diferentes erro-padrão, a única estimativa que se mantém constante para todos os modelos é a mediana. Para as outras estimativas, os valores tornam-se rapidamente muito diferentes.

Estimativa / Erro-Padrão	0,136	0,25	0,5	0,75
Moda	944.013,56	903.396,57	748.942,06	547.937,72
Dif. em relação à Mediana	-1,84%	-6,06%	-22,12%	-43,02%
Mediana	961.660,64	961.660,64	961.660,64	961.660,64
Média	970.607,51	992.187,03	1.089.704,27	1.273.993,36
Dif. em relação à Mediana	+0,93%	+3,17%	+13,31%	+32,48%

Tabela 2: Estimativas Moda, Média e Mediana para diferentes valores de erro-padrão

4 I CONCLUSÕES E RECOMENDAÇÕES

Conforme se discute em DROUBI et al. (2018), a transformação da variável dependente na regressão linear tem por objetivo tentar remover a heteroscedasticidade do modelo, o que acarreta em distorções, haja vista que a regressão assim obtida é valida apenas para a variável transformada, já que na escala original esta equação de regressão difere da equação da escala logarítmica, pela desigualdade de Jensen.

Conforme apresentado em ZONATO et al. (2018), existem outras maneiras de se contornar o problema da heteroscedasticade, através do uso de métodos que computem erros heteroscedásticos-consistentes, como o método de Eicker-White, ou através da utilização de regressão ponderada.

Recomenda-se, desta maneira, que seja evitada a utilização de transformações nos modelos de regressão, sempre que possível e, em caso de heteroscedasticidade, utilizar os métodos supra-citados. No entanto, se a transformação da variável dependente for necessária, recomenda- se especial atenção à heteroscedasticidade, fazendo uso de métodos como o de Box-Cox para encontrar a transformação que melhor estabiliza a variância do modelo.

Em caso de transformação da variável dependente pela função logaritmo natural, deve ser escolhida a estimativa adequada. Como foi visto na secção 2.10, o método clássico de regressão linear é uma minimização do erro médio quadrático de predição e a função de regressão $m_{Y;X}$ é uma equação para a média da população *Y* dado *X*, seja ela uma função de outra variável ou não.

Ora, claro está, de acordo com todos os trabalhos citados, inclusive GIANNAKOS; LEÃO (1996), que o valor esperado da variável é a média. A regressão linear com o método dos mínimos quadrados é uma regressão para a média. Isto posto, como então avaliar o valor da variável original? Porque na área de avaliações não há interesse na previsão da variável $W = \ln(Y)$, mas sim na variável *Y*, ou seja, existe interesse nos valores previstos para a variável original, não nos valores da variável transformada. Está claro que deve-se proceder a retransformação da variável *W* na variável original, mas para isso, qual estimativa utilizar?

Conforme mostrado, matematicamente as três estimativas são válidas. No âmbito da Engenharia de Avaliações, no entanto, para a determinação do valor de um imóvel em específico, entende-se que não seria ideal que se utilizasse a avaliação pela média ou pela moda da variável lognormal, haja vista que, conforme demonstrado, os seus valores podem variar bastante de um modelo para outro, a depender do erro-padrão.

Assim, poder-se-ia imaginar hipoteticamente que, dois avaliadores, de maneira independente, ao estudar um determinado mercado para a avaliação de um imóvel cheguem a modelos de regressão semelhantes, com transformação da variável dependente pela função logaritmo natural, obtendo-se valores semelhantes dos coeficientes de regressão. Porém, a depender de suas amostras, um dos modelos pode ter um erro-padrão diferente do outro. Estes dois avaliadores, ao avaliarem o imóvel em pauta pela mediana da variável lognormal, chegariam ao mesmo resultado. Porém, se os mesmos adotarem a média ou a moda da variável lognormal, estes valores podem ser significativamente diferentes. A situação ainda se agravaria caso um dos avaliadores adotasse a avaliação pela média e o outro a avaliação pela moda.

Em vários campos, a mediana tem sido adotada como melhor estimativa, por sua propriedade de estar menos vulnerável a presença de *outliers*, o que não ocorre com a média.

REFERÊNCIAS

ABNT. **NBR 14653-2: Avaliação de bens – parte 2: Imóveis urbanos**. Rio de Janeiro: Associação Brasileira de Normas Técnicas, 2011.

BENNETT, H. Lecture note 4: Expectations (moments)., 2006. MIT. Disponível em: <https://tinyurl.com/yayljdpq>..

DROUBI, L. F. P.; ZONATO, W.; HOCHHEIM, N. Distribuição Lognormal: Propriedades e aplicações na engenharia de avaliações. In: 13º Congresso Brasileiro de Cadastro Técnico Multifinalitário e Gestão Territorial. **Anais…**, 2018. Florianópolis: COBRAC. Disponível em: <http://droubi.me/cobrac2018>..

DUAN, N. Smearing estimate: A nonparametric retransformation method. **Journal of the American Statistical Association**, v. 78, n. 383, p. 605–610, 1983. Taylor & Francis. Disponível em: <http://www.tandfonline.com/doi/abs/10.1080/01621459.1983.10478017>..

GIANNAKOS, I. B. D. S.; LEÃO, M. L. Crítica à avaliação pela moda da distribuição log- normal. In: VIII Congresso Brasileiro de Avaliações e Perícias. **Anais…**. p.267–278, 1996. Flori- anópolis: COBREAP. 2015.

HOCHHEIM, N. **Engenharia de avaliações - módulo básico**. Florianópolis: IBAPE - SC, 2015

LIMPERT, E.; A. STAHEL, W.; ABBT, M. Log-normal distributions across the sciences: Keys and clues., v. 51, p. 341–352, 2001.

MANNING, W. G.; MULLAHY, J. **Estimating log models: To transform or not to transform?** Working Paper, National Bureau of Economic Research, 1999.

MATLOFF, N. **Statistical regression and classification: From linear models to ma- chine learning**. Boca Raton, Florida: Chapman & Hall, 2017.

MATLOFF, N. S. **From Algorithms to Z-Scores: Probabilistic and statistical modeling in computer science**. Davis, California: Orange Grove Books, 2009.

SHEN, H.; ZHU, Z. Efficient mean estimation in log-normal linear models. **Journal of Statistical Planning and Inference**, v. 138, p. 552–567, 2008. Elsevier. Disponível em: <https://www.unc.edu/~haipeng/publication/emplnM1.pdf>..

WASSERMAN, L. **All of statistics: A concise course in statistical inference**. Springer Publishing Company, Incorporated, 2010.

ZONATO, W.; DROUBI, L. F. P.; HOCHHEIM, N. Pressupostos clássicos dos modelos de regressão linear e suas implicações sobre as avaliações em massa. In: 13º Congresso Brasileiro de Cadastro Técnico Multifinalitário e Gestão Territorial. **Anais...**, 2018. COBRAC. Disponível em: <http://droubi.me/cobrac2018>..

SOBRE O ORGANIZADOR

CLEBERTON CORREIA SANTOS- Graduado em Tecnologia em Agroecologia, mestre e doutor em Agronomia (Produção Vegetal). Tem experiência nas seguintes áreas: agricultura familiar, indicadores de sustentabilidade de agroecossistemas, uso e manejo de resíduos orgânicos, propagação de plantas, manejo e tratos culturais em horticultura geral, plantas medicinais exóticas e nativas, respostas morfofisiológicas de plantas ao estresse ambiental, nutrição de plantas e planejamento e análises de experimentos agropecuários.

(E-mail: cleber_frs@yahoo.com.br) – ORCID: 0000-0001-6741-2622

ÍNDICE REMISSIVO

Agência Brasileira do ISBN
ISBN 978-85-7247-621-8
9 788572 476218