

Distribuição Log-Normal

Propriedades e aplicações

*Luiz Fernando Palin Droubi**

Norberto Hochheim†

Willian Zonato‡

22/06/2018

Resumo

Pretende-se com este artigo detalhar o motivo pelo qual a transformação de variável dependente pela função logaritmo é frequentemente adequada na área de avaliação de imóveis. Um procedimento muito comum nesta área é a adoção de transformações para a obtenção de um “melhor” modelo de regressão. A mais usual e preferida de muitos avaliadores é a função logaritmo, especialmente para a variável dependente. Muitas vezes esta transformação é adequada e percebe-se uma notória melhora no ajuste do modelo. Outras vezes, esta transformação pode não ser adequada. Apesar do modelo aparentar-se melhor ajustado, problemas podem ocorrer quanto às verificações das hipóteses clássicas da regressão, as quais nem sempre os avaliadores estão tão atentos quanto estão com as verificações dos intervalos de confiança e níveis de significância. No entanto, o avaliador que assim procede estará verificando intervalos de confiança e níveis de significâncias incorretos, haja vista que a hipótese da heteroscedasticidade implica na incorreção destas inferências. Entendemos que a melhor maneira para apresentar aos avaliadores a importância de criteriosas escolhas de transformações seja através da análise do histograma da variável original e transformada. Normalmente, uma boa escolha de transformação leva à uma distribuição aproximadamente normal. Quando a variável dependente apresenta distribuição lognormal, esta transformação é a transformação logarítmica. Desta maneira, demonstramos as características básicas desta distribuição, sua formulação, características além do seu relacionamento com a distribuição normal. Por fim, demonstramos as implicações da adoção da transformação da variável dependente e abordamos o problema da retransformação da variável dependente à sua escala original.

1 INTRODUÇÃO

A transformação de variáveis é um procedimento comum na Engenharia de Avaliações. No entanto, a transformação dos dados por vezes é realizada sem uma análise profunda do comportamento das variáveis. A *Food and Drug Administration* (FDA), órgão federal dos EUA que atua no controle da comercialização de alimentos e medicamentos no país, recomenda:

A transformação desnecessária de dados deve ser evitada. Caso tenha sido realizada transformação de dados, uma justificativa para a escolha da transformação junto com a interpretação das estimativas dos efeitos do tratamento com base nos dados transformados deve ser fornecida. (FDA, 1988 apud KEENE (1985))

No entanto, a transformação logarítmica é especial, por uma série de aspectos, como pode ser visto em KEENE (1985).

A distribuição lognormal apresenta diversas aplicações práticas. É comum, na área de avaliação de imóveis, mas não apenas¹, nos depararmos com dados que seguem esta distribuição. Neste artigo pretendemos demonstrar as principais características da distribuição lognormal, sua relação com a distribuição normal de Gauss, assim como debatemos a melhor maneira de se lidar com dados lognormais.

*SPU/SC, luiz.droubi@planejamento.gov.br

†UFSC, hochheim@gmail.com

‡SPU/SC, willian.zonato@planejamento.gov.br

¹Dados estritamente positivos, como valores em moeda, altura, peso, etc, normalmente seguem a distribuição lognormal.

2 REVISÃO BIBLIOGRÁFICA

2.1 Formulação

A formulação da distribuição lognormal para os parâmetros μ e σ pode ser vista abaixo (FARIAS)

$$f(x; \mu, \sigma) = \begin{cases} \frac{1}{x\sigma\sqrt{2\pi}} \exp\left(-\frac{(\log(x)-\mu)^2}{2\sigma^2}\right) & \forall x > 0 \\ 0 & \text{se } x = 0 \end{cases}$$

2.2 Propriedades

2.2.1 Valor Esperado e Variância

O valor Esperado \mathbb{E} de uma variável aleatória com distribuição lognormal X é (FARIAS):

$$\mathbb{E}(X) = \exp\left(\mu + \frac{\sigma^2}{2}\right)$$

E sua variância é:

$$\text{Var}(X) = \exp(2\mu + \sigma^2)(\exp(\sigma^2) - 1)$$

2.2.2 Medidas de Tendência Central

A figura 1 mostra a posição das medidas de tendência central (moda, média e mediana) para um variável aleatória de distribuição log-normal.

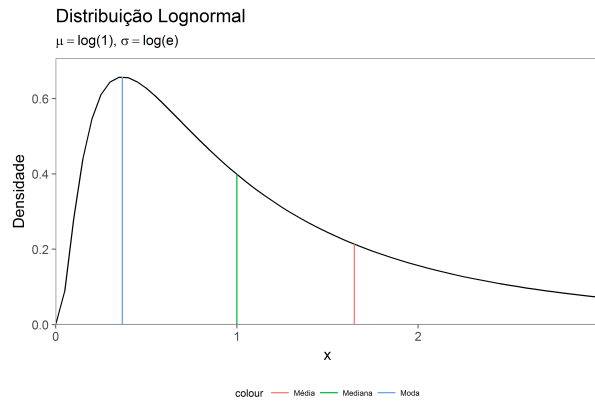


Figura 1: Ilustração das posições de medidas de tendência central numa distribuição lognormal.

2.2.3 Efeito das variações do desvio-padrão na forma da distribuição

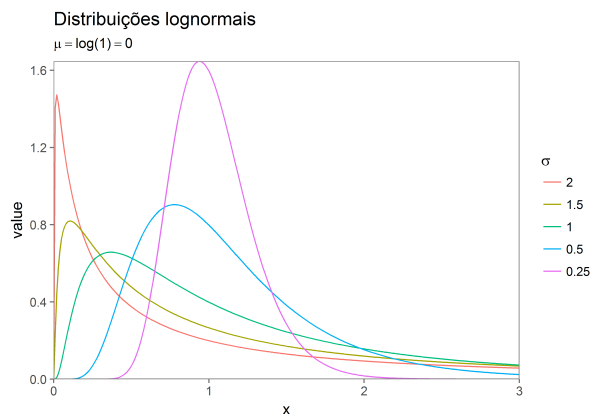


Figura 2: Distribuição lognormal com $\mu = 0$ e diversos valores de σ

2.2.4 Relação com a distribuição normal

Lembrando que a função densidade de probabilidade de uma variável aleatória com distribuição normal é dada por:

$$f(t) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2} \frac{(t-\mu)^2}{\sigma^2}}$$

E que para a distribuição normal-padrão ($N(0, 1)$) a função densidade de probabilidade torna-se:

$$\varphi(t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2} t^2}$$

Seja X uma variável aleatória de distribuição normal padronizada ($X \sim N(0, 1)$), f_X a função densidade de probabilidade e $Y = e^X$. Então (F_Y) é igual a:

$$F_Y(y) = \mathbb{P}(e^X \leq y) = \mathbb{P}(X \leq \ln(Y)) = \int_{-\infty}^{\ln(y)} f_X(x) dx = \int_{-\infty}^{\ln(y)} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx$$

o que equivale a:

$$F_Y(y) = \int_0^y \frac{1}{x} \frac{1}{\sqrt{2\pi}} e^{-\ln(x)^2/2} dx$$

Ou seja, a distribuição de uma variável $Y = e^X$, em que $X \sim N(0, 1)$ é equivalente a distribuição de uma variável lognormal com parâmetros $\mu = 0$ e $\sigma = 1$.

A figura 3 ilustra este fato.

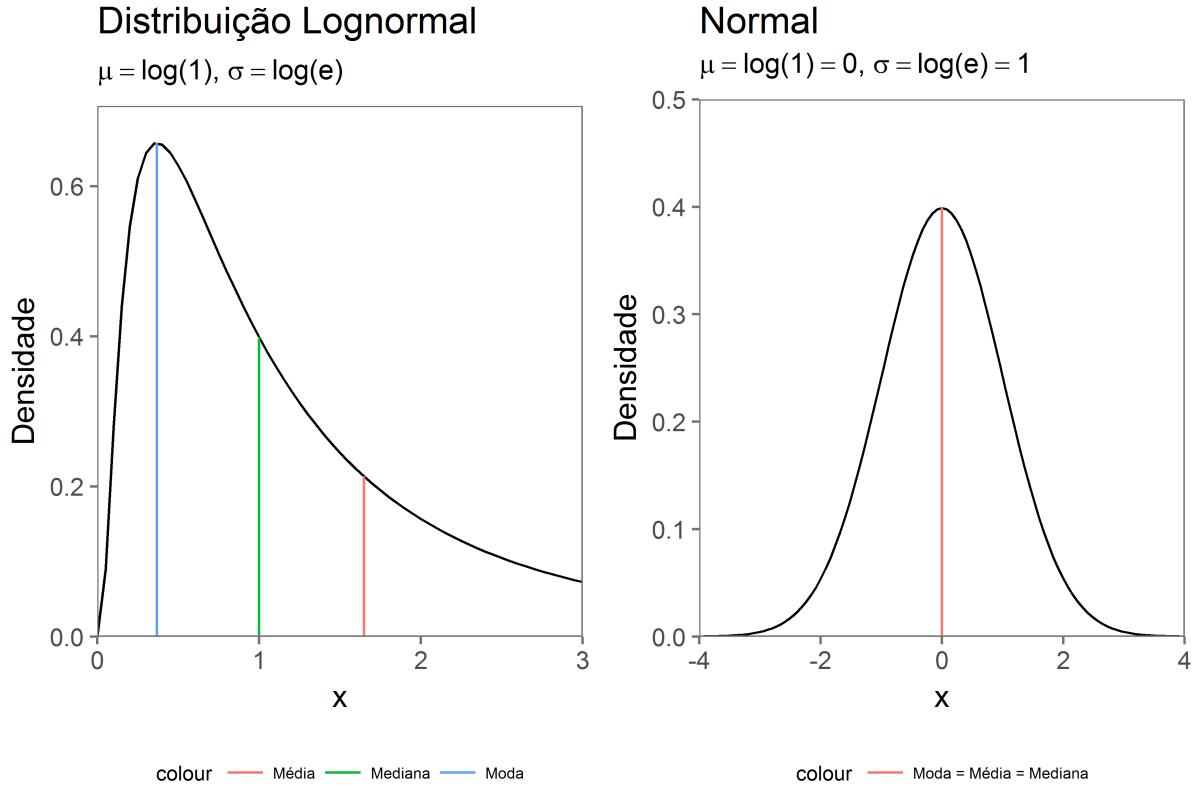


Figura 3: Comparação entre distribuições normal e lognormal padronizadas.

2.2.5 Analogia com o Teorema do Limite Central

Assim como o resultado da soma de diversas variáveis independentes com distribuições quaisquer resulta numa variável aleatória de distribuição normal (Teorema do Limite Central), o produto de diversas variáveis aleatórias resulta numa distribuição lognormal.

2.2.6 Transformação de variável e Homoscedasticidade

De acordo com Matloff (2017, p. 138), se uma variável aleatória W é aproximadamente normal, com baixo coeficiente de variação ($CV = \sigma/\mu$), e $g(W)$ é uma função suave, então a nova variável também será aproximadamente normal, com média $g(EW)$ e variância:

$$[g'(EW)]^2 \text{Var}(W)$$

Assumindo que os erros de uma função de regressão sejam heteroscedásticos, seguindo uma função conhecida $\sigma(t) = \mu(t)$, se aplicarmos a função logaritmo natural à variável dependente, segundo a equação acima, teremos²:

$$\frac{1}{\mu^2(t)} \mu^2(t) = 1$$

Ou seja, o uso da transformação logaritmo natural, para este caso em particular, conduz à homoscedasticidade do modelo.

²Lembrando que a derivada da função logaritmo natural é $\frac{d}{dt} \ln t = \frac{1}{t}$ e que $\text{Var}(W) = \sigma^2(W)$

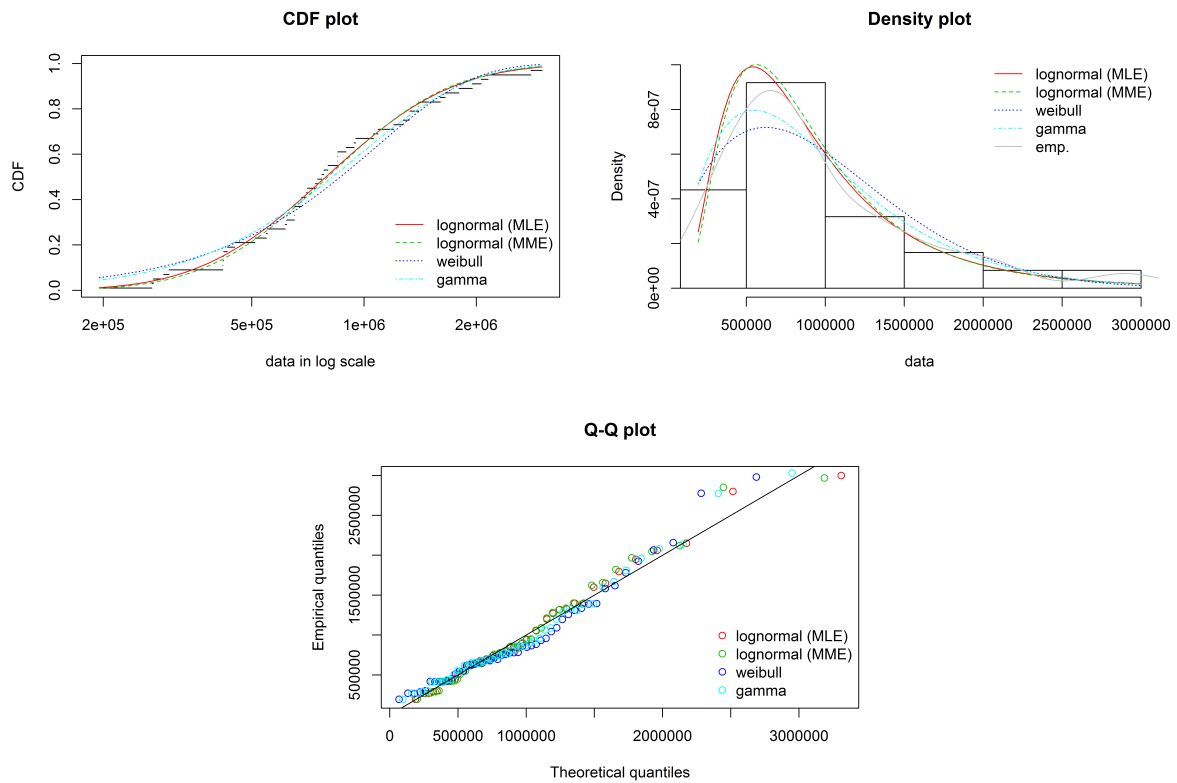
De acordo com Matloff (2017, p. 138), ainda, se $\sigma(t) = \sqrt{\mu(t)}$, a transformação raiz-quadrada é que traria de volta a homoscedasticidade.³

3 EXEMPLO

3.1 Dados

Os dados utilizados aqui são oriundos de Hochheim (2015, pp. 21–22) e são reproduzidos no ANEXO I.

3.2 Ajuste de distribuições aos dados



3.3 Gráficos

As figuras 4 e 5 mostram que os valores observados para a variável **valor** do conjunto de dados mencionados acima (HOCHHEIM, 2015, pp. 21–22) apresentam distribuição aproximadamente lognormal, com parâmetros $\mu = \ln(\text{valor})$

³ $\frac{d}{dt} \sqrt{t} = \frac{0,5}{\sqrt{t}} \rightarrow \text{Var}(\sqrt{W}) = \left(\frac{0,5}{\sqrt{t}} \right)^2 (\sqrt{t})^2 = 0,25$

a. Densidade

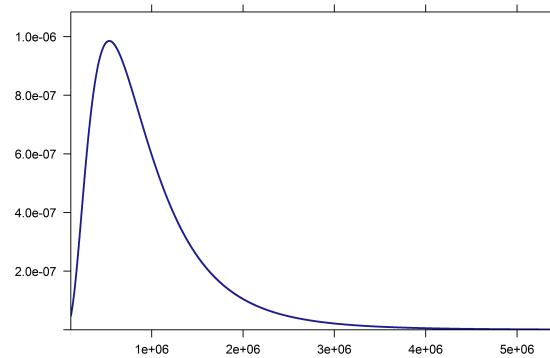


Figura 4: Função densidade de probabilidade com parâmetros obtidos dos dados da variável **valor**

b. Histograma com densidade superposta

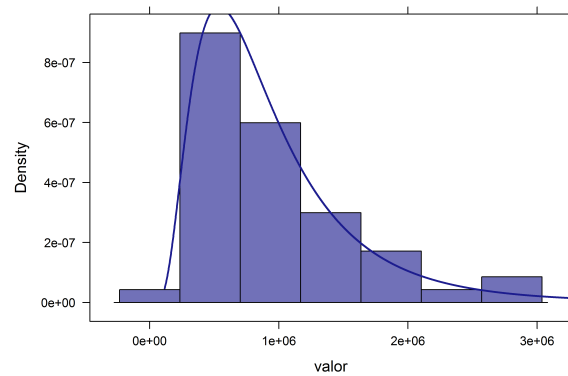


Figura 5: Histograma das variável **valor** com função densidade de probabilidade superposta.

c. Cumulativa

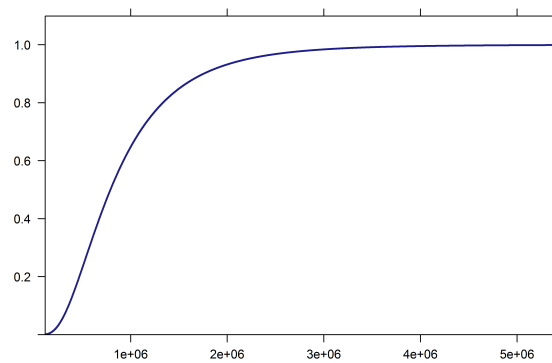


Figura 6: Função cumulativa de densidade de probabilidade com parâmetros obtidos dos dados da variável **valor**

d. Distribuição da variável $\ln(valor)$

A figura 7

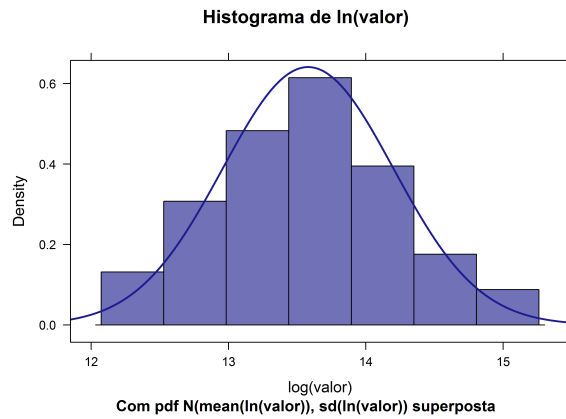


Figura 7: Histograma com função densidade de probabilidade normal superposta

3.4 Modelos

Detectando-se a presença de variável resposta com distribuição lognormal, pode-se proceder da seguinte maneira:

3.4.1 Modelo linear com a variável resposta transformada

É fácil mostrar que o modelo linear com a variável resposta logaritmizada, ou seja, com distribuição normal, é melhor ajustado que o modelo linear de uma variável resposta lognormal. Além disso, o modelo linear, sem transformação, é heteroscedástico.

```
##
## studentized Breusch-Pagan test
##
## data: fit
## BP = 17.882753, df = 7, p-value = 0.01251035
```

A função máxima verossimilhança de Box-Cox também vai apresentar como transformação ótima a transformação logarítmica, como demonstra a figura 8

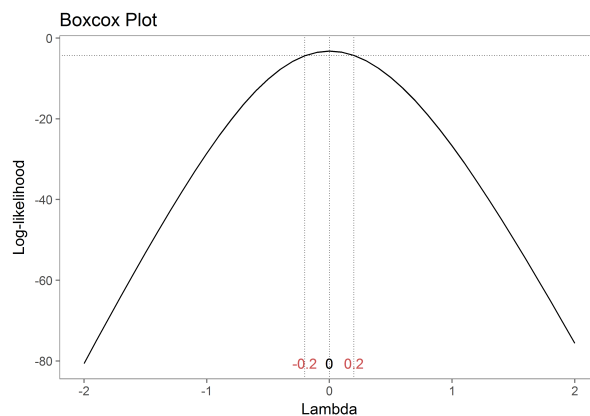


Figura 8: Gráfico da função verossimilhança de Box-Cox

Na tabela 1 é possível comparar os modelos com e sem a transformação da variável resposta. Porém, como o modelo sem transformação é heteroscedástico, os intervalos de confiança dos regressores e os p-valores mostrados na tabela são inválidos, pois deve-se calcular os erros robustos antes de computá-los, o que será visto na próxima seção.

Tabela 1: Comparação entre modelos com e sem transformação da variável resposta

	<i>Dependent variable:</i>	
	valor (1)	log(valor) (2)
area_total	2,893.178 (2,065.405, 3,720.951) t = 6.850 p = 0.00000***	0.002 (0.001, 0.002) t = 4.886 p = 0.00002***
quartos	73,524.375 (−34,814.143, 181,862.894) t = 1.330 p = 0.191	0.169 (0.084, 0.255) t = 3.870 p = 0.0004***
suítes	111,000.591 (8,045.131, 213,956.052) t = 2.113 p = 0.041**	0.088 (0.007, 0.170) t = 2.121 p = 0.040**
garagens	148,427.448 (49,657.102, 247,197.795) t = 2.945 p = 0.006***	0.175 (0.097, 0.253) t = 4.394 p = 0.0001***
dist_b_mar	−223.217 (−434.862, −11.571) t = −2.067 p = 0.045**	−0.0003 (−0.0004, −0.0001) t = −3.215 p = 0.003***
padraomedio	−146,549.393 (−354,850.457, 61,751.672) t = −1.379 p = 0.176	0.268 (0.103, 0.433) t = 3.190 p = 0.003***
padraoalto	−56,064.550 (−264,003.525, 151,874.426) t = −0.528 p = 0.600	0.334 (0.169, 0.498) t = 3.975 p = 0.0003***
Constant	33,953.788 (−267,469.800, 335,377.375) t = 0.221 p = 0.827	12.315 (12.076, 12.553) t = 101.170 p = 0.000***
Observations	50	50
R ²	0.906	0.940
Adjusted R ²	0.890	0.930
Akaike Inf. Crit.	1,375.659	−29.275
Residual Std. Error (df = 42)	207,903.003	0.165
F Statistic (df = 7; 42)	57.731***	94.063***

Note:

*p<0.1; **p<0.05; ***p<0.01

3.5 Retransformação de variáveis

O problema da transformação da variável resposta no logaritmo da variável resposta original, é que devemos estudar como proceder na retransformação da variável, para efetuar a avaliação do imóvel.

3.5.1 A desigualdade de Jensen

Segundo Matloff (2017, p. 142), a desigualdade de Jensen (aplicada à estatística) se traduz na seguinte expressão, válida para funções convexas:

$$\mathbb{E}[h(V)] \geq h(\mathbb{E}[V])$$

Isto aplicado no caso da transformação logarítmica, que é uma função côncava, se reduz à expressão abaixo (MATLOFF, 2017, p. 142):

$$\mathbb{E}[\ln Y|X = t] \leq \ln(E[Y|X = t])$$

Para Matloff, então, como a igualdade só irá acontecer em poucos casos especiais, a função de regressão de $\ln(Y)$ será quase sempre menor do que o logaritmo natural da função de regressão de Y , de tal forma que a suposição que dado uma variável aleatória Y tal que assumimos que $E(Y|X = t) = e^{\beta_0 + \beta_1 t}$, não podemos concluir de imediato que um modelo linear razoável seria da forma $E(\ln Y|X = t) = \beta_0 + \beta_1 t$, pois, pela desigualdade de Jensen, se temos dados significantemente heteroscedásticos da variável original (Y), a discrepância entre os dois lados da desigualdade acima poderia variar bastante com t , potencialmente produzindo uma grande distorção à forma da curva de regressão (MATLOFF, 2017, p. 143). Segundo Becker (2012, p. 4), a desigualdade de Jensen pode ser transformada numa igualdade do tipo:

$$\mathbb{E}[f(X)] = f(\mathbb{E}[X]) + \Delta$$

E, de acordo com o mesmo (2012, p. 5), o valor de Δ , chamado de Defeito de Holder, é proporcional à variância da variável aleatória X , tal que se $f : [a, b] \rightarrow \mathbb{R}$ é duas vezes continuamente diferenciável e existem limites finitos m e M tais que $0 \leq m \leq f''(x) \leq M$ para todo $x \in [a, b]$, então existe um valor $\mu \in [m, M]$ para o qual a fórmula abaixo é válida:

$$\mathbb{E}[f(X)] - f(\mathbb{E}[X]) = \frac{1}{2}\mu\text{Var}(X)$$

Em suma, o valor de Δ é proporcional à variância de X ($\Delta \propto \sigma^2(X)$).

Deste modo, existem na literatura diversos estudos sobre qual seria o “melhor” estimador – paramétrico ou não-paramétrico – para a variável resposta original, quando da ocorrência da transformação da variável pela função logaritmo natural, como pode ser visto em DUAN (1983), MEULENBERG (1965) e SHEN; ZHU (2008).

Entendemos que, na precisão necessária para a área de avaliação de imóveis, é suficiente a adoção do estimador teórico, apesar do funcionamento dos estimadores não-paramétricos ter sido demonstrado mais eficiente do que ele.

$$\mathbb{E}(Y|X) = \exp(\beta_0 + \beta_1 X + 0.5\sigma^2)$$

3.5.2 Modelo linear com posterior correção da heteroscedasticidade

3.5.2.1 Erros-padrão robustos

a. Coeficientes

b. Teste F

Na tabela 2, o resultado para o teste F para o modelo robusto está errado. O teste deve ser feito com a consideração dos erros robustos:

Tabela 2: Comparação entre modelos com e sem erros robustos

	<i>Dependent variable:</i>	
	valor	
	default (1)	robust (2)
area_total	2,893.178 (2,065.405, 3,720.951) t = 6.850 p = 0.00000***	2,893.178 (1,712.055, 4,074.302) t = 4.801 p = 0.00001***
quartos	73,524.375 (-34,814.143, 181,862.894) t = 1.330 p = 0.191	73,524.375 (-31,897.755, 178,946.506) t = 1.367 p = 0.172
suites	111,000.591 (8,045.131, 213,956.052) t = 2.113 p = 0.041**	111,000.591 (18,671.249, 203,329.934) t = 2.356 p = 0.019**
garagens	148,427.448 (49,657.102, 247,197.795) t = 2.945 p = 0.006***	148,427.448 (73,906.905, 222,947.992) t = 3.904 p = 0.0001***
dist_b_mar	-223.217 (-434.862, -11.571) t = -2.067 p = 0.045**	-223.217 (-406.588, -39.845) t = -2.386 p = 0.018**
padraomedio	-146,549.393 (-354,850.457, 61,751.672) t = -1.379 p = 0.176	-146,549.393 (-322,293.226, 29,194.441) t = -1.634 p = 0.103
padraoalto	-56,064.550 (-264,003.525, 151,874.426) t = -0.528 p = 0.600	-56,064.550 (-197,646.893, 85,517.794) t = -0.776 p = 0.438
Constant	33,953.788 (-267,469.800, 335,377.375) t = 0.221 p = 0.827	33,953.788 (-191,823.586, 259,731.161) t = 0.295 p = 0.769
Observations	50	50
R ²	0.906	0.906
Adjusted R ²	0.890	0.890
Akaike Inf. Crit.	1,375.659	1,375.659
Residual Std. Error (df = 42)	207,903.003	207,903.003
F Statistic (df = 7; 42)	57.731***	57.731***

Note:

*p<0.1; **p<0.05; ***p<0.01

```
## Wald test
##
## Model 1: valor ~ area_total + quartos + suites + garagens + dist_b_mar +
##      padrao
## Model 2: valor ~ 1
##      Res.Df Df       F      Pr(>F)
## 1         42
## 2         49 -7 33.42548 3.3877e-15 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

3.5.3 Comparação das previsões para os dois modelos

4 CONCLUSÃO

Foi possível demonstrar de maneira gráfica que os dados da variável **valor** apresentados se ajustam bem a uma distribuição lognormal equivalente. Por definição, então, o logaritmo da variável possui distribuição normal.

O valor mais provável para a variável resposta, então, é Valor Esperado da variável. Logo, a retransformação da variável deve ser feita para a média da variável log-normal.

fit	lwr	upr	amplitude	
1.106.966,61	1.012.097,40	1.210.728,40	0,18	
665.565,24	624.364,31	709.484,96	0,13	
889.515,06	839.323,09	942.708,54	0,12	
778.638,69	732.296,73	827.913,31	0,12	
926.373,32	826.896,03	1.037.817,93	0,23	
366.492,31	337.377,96	398.119,10	0,17	
720.258,16	676.775,68	766.534,37	0,12	
679.913,27	635.304,18	727.654,68	0,14	
677.695,98	633.095,02	725.439,04	0,14	
703.407,28	661.531,44	747.933,93	0,12	
855.019,06	805.611,95	907.456,25	0,12	
743.187,12	699.766,15	789.302,39	0,12	
747.511,83	703.804,25	793.933,73	0,12	
479.030,46	425.602,14	539.165,94	0,24	
685.488,30	640.831,29	733.257,28	0,13	
1.339.514,49	1.217.160,39	1.474.168,13	0,19	
1.116.148,08	1.040.974,23	1.196.750,59	0,14	
663.458,78	618.767,51	711.377,93	0,14	
712.383,70	668.349,56	759.319,03	0,13	
431.580,00	389.600,88	478.082,34	0,21	
243.707,09	218.091,05	272.331,87	0,22	
483.061,17	436.055,79	535.133,57	0,21	
629.182,10	579.007,59	683.704,53	0,17	
778.457,88	730.734,27	829.298,28	0,13	
720.258,16	676.775,68	766.534,37	0,12	
385.471,14	354.491,70	419.157,91	0,17	
236.922,70	216.532,40	259.233,11	0,18	
289.533,09	265.323,28	315.951,95	0,17	
233.663,54	209.605,75	260.482,61	0,22	
398.390,85	358.851,53	442.286,72	0,21	
660.321,06	621.395,99	701.684,44	0,12	
465.669,76	433.446,51	500.288,54	0,14	
1.582.956,13	1.443.690,00	1.735.656,62	0,18	
2.532.821,47	2.218.551,51	2.891.609,49	0,27	
970.590,45	911.737,58	1.033.242,29	0,13	
1.219.926,54	1.095.450,47	1.358.546,82	0,22	
1.022.195,60	909.895,02	1.148.356,49	0,23	
1.129.958,73	1.033.156,70	1.235.830,67	0,18	
2.093.864,11	1.905.595,34	2.300.733,44	0,19	
1.938.988,72	1.716.090,88	2.190.838,08	0,24	
1.506.487,83	1.389.422,95	1.633.415,93	0,16	
3.464.954,28	3.132.699,80	3.832.447,71	0,20	
1.178.021,65	1.080.266,41	1.284.622,95	0,17	
695.650,38	633.799,37	763.537,29	0,19	
653.267,76	595.450,52	716.698,95	0,19	
2.284.770,17	2.093.977,72	2.492.946,64	0,17	
1.639.750,58	1.523.582,17	1.764.776,47	0,15	
813.100,32	762.277,24	867.311,91	0,13	
1.587.840,38	1.464.634,52	1.721.410,37	0,16	
802.909,81	750.036,73	859.510,12	0,14	
predicted.value	se	ci.lower	ci.upper	amplitude
1.367.729,29	121.445,88	1.212.090,13	1.523.368,45	0,23
614.058,83	54.198,28	544.600,94	683.516,72	0,23
899.907,02	44.247,13	843.202,04	956.611,99	0,13
801.067,23	58.642,85	725.913,40	876.221,07	0,19
910.856,15	108.475,03	771.839,81	1.049.872,49	0,31
445.349,38	61.718,67	366.253,72	524.445,03	0,36
844.240,07	47.795,77	782.987,32	905.492,82	0,15

ANEXO I

valor	area_total	quartos	suites	garagens	dist_b_mar	padrao
1060000	350.00	3	1	2	720	medio
510000	136.56	3	1	1	665	medio
780000	164.77	3	1	2	415	medio
550000	174.58	3	1	1	320	medio
850000	123.01	3	1	3	895	alto
300000	89.83	2	0	1	645	baixo
750000	174.00	2	1	2	860	alto
650000	123.00	3	1	1	745	alto
620000	121.00	3	1	1	745	alto
740000	109.00	3	1	1	300	medio
770000	170.00	3	1	2	590	medio
680000	141.00	3	1	1	290	medio
850000	174.00	3	1	1	465	medio
420000	105.00	3	1	0	60	baixo
547000	128.00	3	1	1	745	alto
1600000	163.00	4	2	2	90	alto
1320000	230.00	3	1	2	215	alto
615000	108.00	3	1	1	745	alto
705000	174.00	2	1	2	900	alto
418000	85.00	1	0	1	620	alto
270000	71.00	2	0	0	1380	baixo
418000	100.00	1	1	1	620	alto
650000	90.00	2	1	1	215	alto
700000	161.00	2	1	2	500	alto
680000	174.00	2	1	2	860	alto
420000	76.00	2	1	1	700	baixo
195000	48.00	1	0	0	730	baixo
290000	66.00	1	0	1	745	baixo
272000	50.00	1	0	1	1430	baixo
430000	61.00	2	0	1	170	baixo
895000	109.00	3	1	1	530	medio
450000	89.00	2	0	1	745	medio
1950000	393.00	3	1	3	550	alto
2150000	578.00	3	2	3	260	alto
940000	182.00	3	1	2	200	medio
1400000	262.00	4	1	1	60	alto
1090000	205.00	3	0	3	465	medio
1272000	196.00	3	3	2	610	alto
2800000	463.00	3	3	3	590	alto
1796000	273.00	3	3	4	140	medio
1400000	330.00	4	2	2	655	alto
3000000	533.00	4	3	4	427	alto
1200000	221.00	3	3	2	607	alto
800000	220.00	3	1	1	1000	medio
950000	127.00	2	1	1	60	medio
2061000	362.00	3	3	4	310	alto
1326000	315.00	3	3	3	600	alto
850000	151.00	3	1	2	660	medio
1650000	246.00	3	3	3	307	alto
650000	159.72	3	1	1	120	medio

REFERÊNCIAS

- BECKER, R. A. The variance drain and Jensen's inequality. **CAEPR Working Paper**, 2012. Disponível em: <<http://dx.doi.org/10.2139/ssrn.2027471>>..
- 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>>..
- FARIAS, A. M. L. DE. **Métodos estatísticos aplicados à economia II: Variáveis aleatórias contínuas**. Universidade Federal Fluminense,.
- FDA. **Guideline for the format and content of the clinical and statistical sections of new drug applications**. Food and Drug Administration, Public Health Service, US Department of Health and Human Services, 1988.
- HOCHHEIM, N. **Engenharia de avaliações - módulo básico**. Florianópolis: IBAPE - SC, 2015.
- KEENE, O. N. The log transformation is special. **Statistics in Medicine**, v. 14, p. 811–819, 1985.
- MATLOFF, N. **Statistical regression and classification: From linear models to machine learning**. Boca Raton, Florida: Chapman & Hall, 2017.
- MEULENBERG, M. T. G. On the estimation of an exponential function. **Econometrica**, v. 33, n. 4, p. 863–868, 1965. [Wiley, Econometric Society]. Disponível em: <<http://www.jstor.org/stable/1910362>>..
- 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/emphnM1.pdf>>..