O modelo:

$$\log Y = \beta_0 + \beta_1 \log x + \varepsilon \quad \text{onde} \quad \varepsilon \sim N(0, \alpha^2)$$

ou $\log Y \sim \text{Normal}\,(\mu = \beta_0 + \beta_1 \log x \,;\, \alpha^2)$

Em geral, estamos interessados nas propriedades estatísticas de Y e não de $\log Y$:

$$e^{\log Y} = e^{\beta_0 + \beta_1 \log x + \varepsilon}$$

$$Y = e^{\beta_0 + \log x^{\beta_1} + \varepsilon}$$

$$= e^{\beta_0} \times e^{\log x^{\beta_1}} \times e^{\varepsilon}$$

$$= \beta_0^* \cdot x^{\beta_1} \cdot e^{\varepsilon}, \quad \text{onde } \beta_0^* = e^{\beta_0}$$

↳ definitivamente NÃO é um modelo linear

e^{ε} caracteriza uma variável aleatória do tipo log-normal:

$$E(e^{\varepsilon}) = e^{\alpha^2/2} \quad \text{e} \quad Var(e^{\varepsilon}) = (e^{\alpha^2} - 1) \cdot e^{\alpha^2}$$

Avaliando SOMENTE $E(Y)$ e $Var(Y)$:

$$E(Y) = E(\beta_0^* \cdot x^{\beta_1} \cdot e^{\varepsilon})$$

$$= \beta_0^* \cdot x^{\beta_1} \cdot E(e^{\varepsilon})$$

$$= [\beta_0^* \cdot x^{\beta_1}] \cdot e^{\alpha^2/2}$$

$$= (\beta_0^* \cdot e^{\alpha^2/2}) \cdot x^{\beta_1}$$

$$Var(Y) = Var(\beta_0^* \cdot x^{\beta_1} \cdot e^{\varepsilon})$$

$$= (\beta_0^* \cdot x^{\beta_1})^2 \cdot Var(e^{\varepsilon})$$

$$= (\beta_0^* \cdot x^{\beta_1})^2 \cdot (e^{\alpha^2} - 1) \cdot e^{\alpha^2}$$

comparando:

$$\boxed{Var(Y) \propto E(Y)^2}$$

modelo heterocedástico

Transformação de Box-Cox :

$$Y_i^*(\lambda) = \begin{cases} \ln(Y_i) , & \text{se } \lambda = 0 \\ \dfrac{Y_i^\lambda - 1}{\lambda} , & \text{se } \lambda \neq 0 \end{cases}$$

o valor de λ é escolhido (ou estimado), de forma que:

$$Y_i^*(\lambda) \sim \text{Normal}(\mu ; \sigma^2)$$

No caso de um modelo de regressão linear :

$$Y_i^*(\lambda) \sim \text{Normal}(\mu = \beta_0 + \beta_1 x_i ; \sigma^2)$$

que é o mesmo que ajustar o modelo:

$$Y_i^*(\lambda) = \beta_0 + \beta_1 x_i + \varepsilon_i$$

$$\varepsilon_i \underset{iid}{\sim} \text{Normal}(0 ; \sigma^2)$$