



**UNIVERSIDADE ESTADUAL DA PARAÍBA  
CAMPUS I  
CENTRO DE CIÊNCIAS E TECNOLOGIA  
DEPARTAMENTO DE ESTATÍSTICA  
CURSO DE BACHARELADO EM ESTATÍSTICA**

**ELIZANDRA DA SILVA BARBOSA**

**USO DE MODELOS ADITIVOS GENERALIZADOS PARA  
LOCAÇÃO, ESCALA E FORMA APLICADOS A DADOS  
SOCIOECONÔMICOS**

**CAMPINA GRANDE - PB**

**2022**

ELIZANDRA DA SILVA BARBOSA

**USO DE MODELOS ADITIVOS GENERALIZADOS PARA LOCAÇÃO, ESCALA E  
FORMA APLICADOS A DADOS SOCIOECONÔMICOS**

Trabalho de Conclusão de Curso apresentado ao curso de Bacharelado em Estatística do Departamento de Estatística do Centro de Ciências e Tecnologia da Universidade Estadual da Paraíba, como requisito parcial à obtenção do título de bacharel em Estatística.

**Orientador:** Profa. Dra. Ana Patrícia Bastos Peixoto

**Coorientador:** Prof. Dr. Tiago Almeida de Oliveira

**CAMPINA GRANDE - PB**

**2022**

É expressamente proibido a comercialização deste documento, tanto na forma impressa como eletrônica. Sua reprodução total ou parcial é permitida exclusivamente para fins acadêmicos e científicos, desde que na reprodução figure a identificação do autor, título, instituição e ano do trabalho.

B238u Barbosa, Elizandra da Silva.

Uso de modelos aditivos generalizados para locação, escala e forma aplicados a dados socioeconômicos [manuscrito] / Elizandra da Silva Barbosa. - 2022.

38 p. : il. colorido.

Digitado.

Trabalho de Conclusão de Curso (Graduação em Estatística) - Universidade Estadual da Paraíba, Centro de Ciências e Tecnologia, 2022.

"Orientação : Profa. Dra. Ana Patrícia Bastos Peixoto ,  
Coordenação do Curso de Estatística - CCT."

"Coorientação: Prof. Dr. Tiago Almeida de Oliveira ,  
Coordenação do Curso de Estatística - CCT."

1. Expectativa de vida. 2. Modelos flexíveis. 3.  
Distribuições. 4. Estatística. 5. Modelos aditivos generalizados.  
I. Título

21. ed. CDD 310

ELIZANDRA DA SILVA BARBOSA

USO DE MODELOS ADITIVOS GENERALIZADOS PARA LOCAÇÃO, ESCALA E FORMA  
APLICADOS A DADOS SOCIOECONÔMICOS

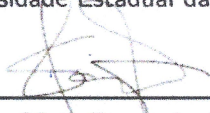
Trabalho de Conclusão de Curso apresentado ao curso de Bacharelado em Estatística do Departamento de Estatística do Centro de Ciências e Tecnologia da Universidade Estadual da Paraíba, como requisito parcial à obtenção do título de bacharel em Estatística.

Trabalho aprovado em 06 de ABRIL de 2022.

**BANCA EXAMINADORA**

Ana Patrícia Bastos Peixoto

Profa. Dra. Ana Patrícia Bastos Peixoto  
Universidade Estadual da Paraíba (UEPB)

  
Prof. Dr. Sílvio Fernando Alves Xavier Júnior  
Universidade Estadual da Paraíba (UEPB)

M<sup>te</sup> das Vitórias A. Serafim

Profa. Ma. Maria das Vitórias Alexandre  
Serafim  
Universidade Estadual da Paraíba (UEPB)

À Deus em sua infinita bondade.  
Aos meus familiares e amigos.  
Com carinho, Dedico.

## AGRADECIMENTOS

Antes de tudo, quero agradecer a Deus por sua infinita bondade comigo e por estar presente em cada decisão tomada e situação que passei durante todo esse caminho pra chegar até aqui.

Aos meus pais, por ter me educado e ensinado a ser a pessoa que sou hoje, forte e corajosa para ir atrás dos meus sonhos e realizar todos os meus objetivos tanto pessoais como profissionais.

As minhas amigas da Universidade como Raquel, Mayara e principalmente Josefa que está comigo desde o início dessa jornada passando por maus bocados, mas também por vários momentos de alegrias que vivenciamos dentro de todo o processo até chegar aqui.

Aos meus amigos Gilmar, Samuel, Mateus, Wyllian e entre outros, por todo apoio e carinho para comigo em toda a jornada.

À minha amiga Thaylla que mesmo de longe me apoiava a prosseguir e finalizar este sonho.

Aos meus companheiros de ônibus em que a diversão era garantida e que me faziam ter momentos leves em dias difíceis.

À todos os professores por quem passei e adquiri conhecimentos profissionais e pessoais.

À minha orientadora, Profa. Ana Patrícia, por toda paciência, persistência e conselhos dados ao longo da elaboração deste TCC, mas também agradecer por todo o caminho, por histórias compartilhadas e momentos vividos.

Ao meu coorientador, Prof. Tiago Oliveira, por toda ajuda necessária para a conclusão desse TCC, me dando todo apoio e conhecimento necessário.

À quem não mencionei, mas esteve junto, obrigada por toda ajuda.

E não poderia deixar de registrar aqui gratidão à mim por nunca ter desistido mesmo quando as coisas ficaram difíceis e por persistir no caminho para a realização dos meus sonhos.

“Nunca foi sorte sempre foi Deus”  
(Autor desconhecido)

## RESUMO

O presente trabalho aborda uma classe de modelos de regressão que flexibiliza ainda mais as distribuições para as variáveis respostas que são os Modelos Aditivos Generalizados para Localização, Escala e Forma. A análise de regressão é utilizada para modelar diversos fenômenos, e suas modalidades têm sido atendido com precisão de acordo com seu uso, mas existe situações em que apenas eles não conseguem se adequar aos dados, para isso faz-se necessário o ajuste de modelos com pressupostos mais flexíveis na especificação da distribuição da variável resposta, explicando assim o uso dessa modelagem. O estudo da expectativa de vida é de suma importância, pois ele indica a qualidade de vida da população em um determinado lugar, informando o quanto se evolui tanto economicamente como socialmente. O objetivo deste estudo foi modelar a expectativa de vida, por meio do ajuste de Modelos Aditivos Generalizados para Localização, Escala e forma, afim de mostrar a capacidade dessa classe de regressão em modelar os dados. Considerando o parâmetro com efeito polinomial de grau dois na variável independente (PIB per capita) sob a modelagem, o modelo de distribuição Exponencial Potência de Box-Cox apresentou os melhores resultados.

**Palavras-chaves:** Expectativa de vida. Modelos flexíveis. Distribuições. Estatística.



## **ABSTRACT**

The present work approaches a class of regression models that make even more flexible the distributions for the response variables that are the Generalized Additive Models for Location, Scale and Shape. Regression analysis is used to model several phenomena, and its modalities have been met with precision according to its use, but there are situations in which only they cannot adapt to the data, for this it is necessary to adjust models with more flexible assumptions in specifying the distribution of the response variable, thus explaining the use of this model. The study of life expectancy is of paramount importance, as it indicates the quality of life of the population in a given place, informing how much it evolves both economically and socially. The objective of this study was to model life expectancy, through the adjustment of Generalized Additive Models for Location, Scale and Shape, in order to show the ability of this regression class to model the data. Considering the parameter with degree two polynomial effect on the independent variable (GDP per capita) under the modeling, the Box-Cox Power Exponential distribution model showed the best results.

**Keywords:** Life expectancy. Flexible models. Distributions. Statistic.

## LISTA DE ILUSTRAÇÕES

Figura 1 – Diferentes problemas que podem ser diagnosticados pelo <i>worm plot</i> . . .	24
Figura 2 – Exemplos teóricos da Distribuição Exponencial Potência de Box-Cox . . .	27
Figura 3 – Histograma do PIB per capita . . . . .	29
Figura 4 – Histograma da expectativa de vida . . . . .	29
Figura 5 – Gráfico de dispersão da expectativa de vida ao longo do tempo . . . . .	30
Figura 6 – Gráfico de resíduos para Expectativa de Vida . . . . .	30
Figura 7 – Gráficos residuais e <i>Worm plot</i> da Distribuição Normal . . . . .	32
Figura 8 – Gráficos residuais e <i>Worm plot</i> da Distribuição Normal + $\sigma$ . . . . .	32
Figura 9 – Gráficos residuais e <i>Worm plot</i> da Distribuição Gama Generalizada . . .	33
Figura 10 – Gráficos residuais e <i>Worm plot</i> da Distribuição Box Cox t . . . . .	33
Figura 11 – Gráficos residuais e <i>Worm plot</i> da Distribuição Exponencial Potência de Box-Cox . . . . .	34

## LISTA DE TABELAS

Tabela 1 – Interpretação de diferentes formatos de <i>worm plot</i> . . . . .	23
Tabela 2 – Principais distribuições dos modelos probabilísticos ajustados via pacote <i>gamlss</i> utilizando o Critério de Akaike Generalizados . . . . .	31
Tabela 3 – Descrição dos principais modelos ajustados . . . . .	34
Tabela 4 – Estimativas dos parâmetros para o ajuste da Distribuição Exponencial Potência de Box-Cox . . . . .	35

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b>	<b>11</b>
<b>2</b>	<b>REVISÃO DA LITERATURA</b>	<b>13</b>
<b>2.1</b>	<b>Modelos de Regressão</b>	<b>13</b>
<b>2.1.1</b>	<i>Modelos Lineares Generalizados</i>	<b>14</b>
<b>2.1.2</b>	<i>Modelos Aditivos Generalizados</i>	<b>15</b>
<b>2.1.3</b>	<i>Modelos Aditivos Generalizados para Locação, Escala e Forma</i>	<b>16</b>
<b>2.1.3.1</b>	<i>Estimação</i>	<b>17</b>
<b>2.1.3.2</b>	<i>Inferência</i>	<b>19</b>
<b>2.1.3.3</b>	<i>Tipos de distribuições dentro da família GAMLSS</i>	<b>20</b>
<b>2.1.3.4</b>	<i>Critério de seleção dos modelos</i>	<b>21</b>
<b>2.1.3.5</b>	<i>Análise de resíduos</i>	<b>21</b>
<b>3</b>	<b>APLICAÇÃO</b>	<b>25</b>
<b>3.1</b>	<b>Caracterização do estudo</b>	<b>25</b>
<b>3.2</b>	<b>Distribuições de probabilidade</b>	<b>25</b>
<b>4</b>	<b>RESULTADOS E DISCUSSÕES</b>	<b>29</b>
<b>5</b>	<b>CONCLUSÃO</b>	<b>36</b>
	<b>REFERÊNCIAS</b>	<b>37</b>

## 1 INTRODUÇÃO

Expectativa de vida, conhecida também como esperança de vida, está relacionada com a qualidade de vida que um país possui, já que determinados fatores como educação, economia, desenvolvimento social entre outros que influenciam diretamente. O bem estar de um indivíduo está, diretamente, ligado à sua renda. O desenvolvimento econômico de um país só ocorre a partir de sua capacidade de reduzir sua pobreza, o desemprego e as desigualdades, mesmo que a renda per capita tenha aumentado (SEERS, 1970). Sendo assim, o PIB per capita tornou-se um dos indicadores a serem usados como uma primeira aproximação ao grau de desenvolvimento econômico e qualidade de vida.

No início dos anos 90, o Programa das Nações Unidas para o Desenvolvimento (PNUD) desenvolveu o Índice de Desenvolvimento Humano (IDH), com objetivo de medir o progresso humano e as condições de vida das pessoas, evitando a prática usual, adotada por vários anos, de classificar o nível de bem-estar de países ou regiões pelo tamanho de seu PIB per capita. Dentre os componentes básicos do IDH, encontram-se a educação, longevidade e renda, que permitem comparar determinado agrupamento de indivíduos entre si através do tempo ou diferentes agrupamentos (PNUD, 2012). Pelo fato de incluir a renda per capita, o IDH constitui um indicador melhor de bem-estar, qualidade de vida e desenvolvimento.

No que se refere a técnicas de modelagem, a regressão linear é um estudo que permite entender e quantificar a relação de dependência entre a variável resposta com as variáveis explicativas. Com o objetivo de colher informações sobre determinado comportamento de um fenômeno. Por outro lado, os modelos básicos foram criados com pressupostos rígidos, levando a inadequações do estudo em várias situações. Esta metodologia pressupõe que a variável resposta precisa ter distribuição normal com variância constante e que suas observações sejam independentes entre si, o que nem sempre condiz com a realidade dos dados.

Para tentar resolver este problema, Nelder J. A. e Wedderburn (1972) propuseram os Modelos Lineares Generalizados (MLG) que considera aplicar uma transformação na esperança da variável resposta, não diretamente na variável, mas utilizando a função de ligação para modelar a relação entre  $E(Y)$  e as variáveis explicativas. Esta técnica permite ajustar modelos de regressão para variáveis discretas e contínuas, possibilitando a análise de dados por meio de distribuições que pertençam à família exponencial. Com a intenção de adicionar funções de suavização nos Modelos Lineares Generalizados (MLG), Hastie T. e Tibshirani (1990) propuseram os Modelos Aditivos Generalizados (MAG), melhorando o ajuste e resolvendo problemas de alguns casos.

Entretanto, existem casos que não se ajustam à estas modelagens, por exemplo, as distribuições da família exponencial não se adequam, ou quando os parâmetros da distribuição podem ser melhor explicados se também modelados em função de variáveis

explicativas. Sendo assim, os Modelos Aditivos Generalizados para Locação, Escala e Forma (GAMLSS) foram propostos, por Rigby e Stasinopoulos (2005), permitindo que várias distribuições para a variável resposta seja ajustada na modelagem, independente de qual família ela faça parte. Ou seja, não somente o parâmetro de locação é modelado, mas também todo e qualquer parâmetro da distribuição. Com o objetivo de tornar os pressupostos mais flexíveis, que se une em uma série de métodos pré-existentes. Dentre esses métodos, temos técnicas de modelos lineares, não-lineares, Modelos lineares generalizados (MLG) (NELDER J. A. E WEDDERBURN, 1972), Modelos aditivos generalizados (MAG) (HASTIE T. E TIBSHIRANI, 1990) e modelos mistos com o pressuposto de que as observações sejam independentes.

Portanto, diante dos fatos apresentados, o objetivo do estudo é modelar a expectativa de vida, por meio do ajuste de Modelos Aditivos Generalizados para Locação, Escala e Forma, afim de mostrar o quanto essa classe de modelos é flexível e importante para modelagem dos dados.

## 2 REVISÃO DA LITERATURA

Aqui será apresentada de maneira breve a contextualização dos métodos da análise de regressão básica, MLG e MAG, levando a uma introdutória base para compreender os GAMLSS e entender o seu desenvolvimento. Inserindo também a estimação, inferência, critério de seleção do modelo e a análise de resíduos do GAMLSS.

### 2.1 Modelos de Regressão

A análise de Regressão é uma técnica estatística que tem a finalidade de descrever através de um modelo fenômenos físicos ou biológicos, investigando se existe relação entre variáveis dependentes e independentes, a partir do resultado das observações pode ser aplicada a diversas áreas do conhecimento. Esta técnica foi implementada primeiramente pelo estatístico e antropólogo Britânico, Sir Francis Galton, no século XIX (CASELLA G.; BEERGER, 2010). Existem algumas categorias sendo uma delas o Modelo de Regressão Linear (MRL), conforme Draper N. e Smith (1998) e é expresso por:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \varepsilon_i \quad i = 1, \dots, n,$$

tal que,  $X_i$  representa cada observação da variável explicativa X;  $\beta_0$  representa o parâmetro linear da reta, isto é, o ponto inicial para a variável Y;  $\beta_1$  representa o parâmetro angular da reta, ou seja, o grau que a reta faz com o eixo X e também decide o quanto aumenta ou diminui o valor de Y em relação a X; e  $\varepsilon_i$  representa o erro associado a cada observação em relação a reta de regressão linear.

Para que se possam realizar inferências e conclusões para o uso prático o ajuste do modelo de regressão linear simples exige algumas suposições. Segundo Hoffmann (2006) são elas:

- i. Relação linear entre X e Y;
- ii. Os valores de  $X_i$  são fixos, ou seja, X não é uma variável aleatória;
- iii. A média do erro é nula, logo  $E(\varepsilon_i) = 0$ ;
- iv. A variância do erro, é uma constante igual a  $\sigma^2$ , para todos os valores de X;
- v. Os erros são não correlacionados,  $E(\varepsilon_i \varepsilon_j) = 0$  para todo  $i \neq j$ ;
- vi. Os erros têm distribuição normal, ou seja,  $\varepsilon \sim N(0, \sigma^2)$ .

Usualmente, o modelo de regressão linear múltiplo é qualquer modelo de regressão linear com duas ou mais variáveis independentes com o objetivo de prever uma variável dependente Y, isto é, proporciona estimar a resposta de uma variável, levando em conta os valores de duas ou mais variáveis explicativas (GANZOLA, 2002), sendo expresso por:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_k X_{ki} + \varepsilon_i \quad i = 1, \dots, n.$$

O modelo de Regressão assume que os termos dos erros,  $\varepsilon_i$ , são independentes e identicamente distribuídos seguindo uma normal com média zero e variância constante. O que equivale dizer que:

$$Y_i \sim N(\mu_i; \sigma^2),$$

tal que,

$$\mu_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_k X_{ki}.$$

Estes modelos também podem ser escritos de forma matricial, ou seja,

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}_i,$$

e, assim,  $\mathbf{Y} \sim N(\boldsymbol{\mu}; \mathbf{I}\sigma^2)$  e  $\boldsymbol{\mu} = \mathbf{X}\boldsymbol{\beta}$ , em que  $\mathbf{Y}$  é o vetor das variáveis respostas de dimensão  $n \times 1$ ,  $\mathbf{X}$  é a matriz de valores das variáveis explicativas, também chamada de matriz de delineamento, com dimensão  $n \times (k + 1)$ ,  $\boldsymbol{\beta}$  é o vetor de parâmetros desconhecidos a ser estimado pelos dados, de dimensão  $(k + 1) \times 1$ ,  $\boldsymbol{\varepsilon}$  é o vetor de valores aleatórios ou vetor dos erros, com dimensão  $n \times 1$  em que  $\boldsymbol{\varepsilon} \sim N(\mathbf{0}; \mathbf{I}\sigma^2)$  e  $\boldsymbol{\mu}$  é o vetor de médias, de dimensão  $n \times 1$  (RENCHEA A. C.; SCHAALJE, 2008).

### 2.1.1 Modelos Lineares Generalizados

Quando considera-se que a variável resposta tem distribuição normal, esse pressuposto muito exigente exclui uma gama de situações, como, por exemplo, variáveis respostas binárias, de proporção, de contagem ou ainda categóricas. Apesar de ter sido bastante utilizado no meio acadêmico, e ainda ser utilizado até hoje, os modelos de regressão são inapropriados para uma série de situações.

Sendo assim, Nelder J. A. e Wedderburn (1972) propuseram os Modelos Lineares Generalizados (MLG) que consiste em ter um leque de opções para a distribuição da variável resposta para além da normal, e, assim, tornando-se uma flexível ampliação dos MRL.

Os MLG relacionam a variável resposta com as variáveis explicativas por meio de uma função de ligação, que tem como objetivo garantir que os parâmetros assumam quaisquer valores reais preservando os valores ajustados no domínio da distribuição. Ele agrega vários modelos estatísticos e, principalmente, considera as distribuições da família exponencial. Assim, temos  $\mathbf{Y} \sim F\boldsymbol{\varepsilon}(\boldsymbol{\mu}, \boldsymbol{\phi})$ , em que  $\boldsymbol{\mu}$  e  $\boldsymbol{\phi}$  são vetores dos parâmetros de localização e escala da distribuição, respectivamente, sendo que  $\boldsymbol{\phi}$  é constante. Distribuições como a normal, Poisson, binomial, binomial negativa, gama, beta normal inversa, são



exemplos de distribuições que pertencem à família exponencial e podem ser usadas no ajuste de MLG. No ajuste de MLG envolve três componentes, são eles:

- i. Sendo  $\boldsymbol{\mu}$  a média da distribuição, dispões de,  $E(\mathbf{Y})=\boldsymbol{\mu}$ ;
- ii. O componente sistemático é formado pelas variáveis explicativas  $\mathbf{X}$  e os parâmetros do modelo ( $\boldsymbol{\beta}$ ). O denominado preditor linear ( $\boldsymbol{\eta}$ ) e  $\boldsymbol{\eta}$  é um vetor com dimensão  $n$ , então, temos:  $\boldsymbol{\eta}=\mathbf{X}\boldsymbol{\beta}$ ;
- iii. A função de ligação é o terceiro componente de um MLG, com objetivo de relacionar a média ao preditor linear, assim estabelecendo uma relação entre o componente aleatório e o componente sistemático do modelo, então, dispõe de, ( $g(\boldsymbol{\mu}) = \boldsymbol{\eta}$ ), sendo as funções de ligação ( $g(\cdot)$ ) e devem ser diferenciáveis e monótonas, determinadas segundo Demétrio (2001) como,  $g(E[\mathbf{Y}])=g(\boldsymbol{\mu})=\boldsymbol{\eta}=\mathbf{X}\boldsymbol{\beta}$ .

### 2.1.2 Modelos Aditivos Generalizados

Apesar da flexibilização do pressuposto da normalidade dos modelos de regressão, os MLG ainda são limitados na medida em que considera que a relação da média é linear às variáveis explicativas. Hastie T. e Tibshirani (1990) introduziram a técnica de suavização aos MLG, chamando de Modelos Aditivos Generalizados (MAG). Eles foram pensados como uma alternativa para melhorar o ajuste dos MLG, sendo adicionado uma função de suavização não-paramétrica nas covariáveis de forma a deixar os próprios dados conduzirem a sua relação com o preditor ( $\boldsymbol{\eta}$ ), que muitas vezes, acontecem de forma não-linear. Os MAG podem, eventualmente, resolver problemas de resíduos assimétricos e/ou heterocedásticos, por exemplo. Um MAG pode ser escrito como,

$$\mathbf{Y} \stackrel{ind}{\sim} F_{\varepsilon}(\boldsymbol{\mu}, \boldsymbol{\phi})$$

$$\boldsymbol{\eta} = g(\boldsymbol{\mu}) = \mathbf{X}\boldsymbol{\beta} + s_1(\mathbf{x}_1) + \dots + s_J(\mathbf{x}_J),$$

tal que,  $s_j$  é uma função de suavização não-paramétrica aplicada a covariável  $\mathbf{x}_j$  para  $j = 1, \dots, J$  (STASINOPOULOS et al., 2017). Observa-se que não são todas as covariáveis que precisam de funções de suavização, pois as mesmas têm que ser contínuas caso não sejam, não se aplica. Podemos aqui citar algumas funções de suavização como P-splines, splines cúbicas, loess e redes neurais. A spline é uma curva definida matematicamente por dois ou mais pontos de controle e é dessa que os dados estabelecem a relação entre  $\boldsymbol{\eta}=g\boldsymbol{\mu}$  e as variáveis explicativas, que é o conceito principal para fazer uso de uma função de suavização (EILERS P. H.; MARX, 2015).

Entretanto, ao utilizarmos um MAG, devemos ter cuidado ao realizar a análise dos modelos, porque os erros padrões são inválidos na medida que só podem ser calculados em cima da covariável em si, pressupondo uma relação linear, não se importando com o

efeito das funções de suavização como um todo. Um dos aspectos do ajuste de modelos semi-paramétricos é não ser facilmente descritos em uma forma matemática, mas, eles permitem ser expostos de forma gráfica, que assim pode nos remeter várias informações sobre a adequação do ajuste e dos resíduos.

### 2.1.3 Modelos Aditivos Generalizados para Localização, Escala e Forma

Os modelos MLG e MAG ainda são limitados para alguns estudos, já que apenas o parâmetro de localização (média) das distribuições é modelado, e também, a distribuição precisa pertencer à família exponencial. E existem algumas situações em que poderemos requerer mais flexibilidade para a distribuição da variável resposta.

Sendo assim, Rigby e Stasinopoulos (2005) propuseram uma classe de modelos de regressão denominada GAMLSS - Modelos Aditivos Generalizados para Localização, Escala e Forma com a intenção de permitir o ajuste de modelos que aceitem qualquer distribuição para a variável resposta, independente dela pertencer à alguma família de distribuições. Além do parâmetro (média) ser modelada, todos os outros parâmetros de localização, escala e forma da distribuição da variável resposta são modelados. Com isso, todos os parâmetros são modelados em função das variáveis explicativas e, além disso, os preditores também podem incorporar funções não-paramétricas de suavização, efeitos aleatórios, ou outros termos aditivos. Os GAMLSS possuem como casos especiais os modelos MRL, MLG e MAG, mas estes modelos ainda utilizam do pressuposto que as observações da variável resposta  $Y$  sejam independentes entre si.

Um modelo GAMLSS pode ser escrito por,

$$Y \stackrel{ind}{\sim} D(\boldsymbol{\mu}, \boldsymbol{\sigma}, \boldsymbol{\nu}, \boldsymbol{\tau})$$

$$\begin{aligned} \eta_1 &= g_1(\boldsymbol{\mu}) = \mathbf{X}_1\boldsymbol{\beta}_1 + s_{11}(\mathbf{x}_{11}) + \dots + s_1J_1(\mathbf{x}_1J_1) \\ \eta_2 &= g_2(\boldsymbol{\sigma}) = \mathbf{X}_2\boldsymbol{\beta}_2 + s_{21}(\mathbf{x}_{21}) + \dots + s_2J_2(\mathbf{x}_2J_2) \\ \eta_3 &= g_3(\boldsymbol{\nu}) = \mathbf{X}_3\boldsymbol{\beta}_3 + s_{31}(\mathbf{x}_{31}) + \dots + s_3J_3(\mathbf{x}_3J_3) \\ \eta_4 &= g_4(\boldsymbol{\tau}) = \mathbf{X}_4\boldsymbol{\beta}_4 + s_{41}(\mathbf{x}_{41}) + \dots + s_4J_4(\mathbf{x}_4J_4), \end{aligned} \tag{2.1}$$

em que,  $D(\boldsymbol{\mu}, \boldsymbol{\sigma}, \boldsymbol{\nu}, \boldsymbol{\tau})$  é uma distribuição de quatro parâmetros,  $\mu$  normalmente é um parâmetro de localização,  $\sigma$  é constantemente, um parâmetro de escala,  $\nu$  e  $\tau$  são os parâmetros de forma da distribuição, comumente associados à assimetria e curtose, respectivamente.  $\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3, \mathbf{X}_4$  são as matrizes que podem ou não coincidir, ou melhor, o preditor de cada parâmetro da distribuição recebe diferentes variáveis explicativas (RIGBY; STASINOPOULOS, 2005).

O modelo descrito na Equação (2.1) é a definição de GAMLSS com a máxima quantidade de parâmetro implementados no R Core (TEAM, 2018), entretanto, na forma teórica tem-se a possibilidade de admitir uma distribuição de probabilidade com mais de quatro parâmetros, adicionando os preditores na definição. Ainda nos permite definir

modelos intermediários para distribuições com dois ou três parâmetros e ajustar  $\sigma$  e  $\nu$  em função de covariáveis, considerando diminuir a quantidade de preditores.

### 2.1.3.1 Estimação

Os modelos GAMLSS primeiramente determinados na Equação (2.1) tal que  $D(\boldsymbol{\mu}, \boldsymbol{\sigma}, \boldsymbol{\nu}, \boldsymbol{\tau})$  é a distribuição da variável resposta  $\mathbf{Y}$ ,  $\mathbf{X}_k$  são as matrizes que incluem os termos aditivos lineares no modelo,  $\beta_k$  são os parâmetros lineares e  $S_{kj}(x_{kj})$  representam diferentes funções de suavização para diferentes variáveis explicativas ( $x_{kj}$ ) para  $k = 1, 2, 3, 4$  e  $j = 1, \dots, J_k$ . Pondere-se que a maioria das suavizações pode ser escritas na forma (STASINOPOULOS et al., 2017),

$$\mathbf{s}(\mathbf{x}) = \mathbf{Z}\boldsymbol{\gamma},$$

em que, a matriz  $\mathbf{Z}$  base para a suavização e que depende da variável explicativa  $\mathbf{x}$ ;  $\boldsymbol{\gamma}$  é um vetor de parâmetros a serem estimados, submetidos à penalização  $\boldsymbol{\lambda}\boldsymbol{\gamma}^T \mathbf{G}\boldsymbol{\gamma}$ , indicando uma matriz conhecida  $\mathbf{G} = \mathbf{D}^T \mathbf{D}$  em que  $\mathbf{D}$  indica a matriz de diferenças e  $\boldsymbol{\lambda}$  é um vetor ou escalar de hiperparâmetros que ajusta (ou controla) o grau de suavização necessária para o ajuste.

Sendo assim, o modelo da Equação (2.1) pode ser generalizado e reescrito como,

$$\mathbf{Y} | \boldsymbol{\gamma} \stackrel{ind}{\sim} D(\boldsymbol{\mu}, \boldsymbol{\sigma}, \boldsymbol{\nu}, \boldsymbol{\tau})$$

$$\begin{aligned} \boldsymbol{\eta}_1 &= g_1(\boldsymbol{\mu}) = \mathbf{X}_1 \boldsymbol{\beta}_1 + \mathbf{Z}_{11} \boldsymbol{\gamma}_{11} + \dots + \mathbf{Z}_{1J_1} \boldsymbol{\gamma}_{1J_1} \\ \boldsymbol{\eta}_2 &= g_2(\boldsymbol{\sigma}) = \mathbf{X}_2 \boldsymbol{\beta}_2 + \mathbf{Z}_{21} \boldsymbol{\gamma}_{21} + \dots + \mathbf{Z}_{2J_2} \boldsymbol{\gamma}_{2J_2} \\ \boldsymbol{\eta}_3 &= g_3(\boldsymbol{\nu}) = \mathbf{X}_3 \boldsymbol{\beta}_3 + \mathbf{Z}_{31} \boldsymbol{\gamma}_{31} + \dots + \mathbf{Z}_{3J_3} \boldsymbol{\gamma}_{3J_3} \\ \boldsymbol{\eta}_4 &= g_4(\boldsymbol{\tau}) = \mathbf{X}_4 \boldsymbol{\beta}_4 + \mathbf{Z}_{41} \boldsymbol{\gamma}_{41} + \dots + \mathbf{Z}_{4J_4} \boldsymbol{\gamma}_{4J_4}. \end{aligned} \quad (2.2)$$

Relaciona-se como modelos GAMLSS semi-paramétrico o modelo da Equação (2.2), temos os parâmetros de efeitos fixos expressos por

$$\boldsymbol{\beta} = (\beta_1^T, \beta_2^T, \beta_3^T, \beta_4^T)^T,$$

e Stasinopoulos et al. (2017) definem,

$$\boldsymbol{\gamma} = (\gamma_{11}^T, \dots, \gamma_{1J_1}^T, \gamma_{21}^T, \dots, \gamma_{4J_4}^T)^T,$$

os parâmetros de um efeito aleatório assume que todos  $\gamma_{kj}$  são independentes entre si e que,

$$\gamma_{kj} \sim N(0, [\mathbf{G}_{kj}(\boldsymbol{\lambda}_{kj})]^{-1}),$$

tal que,  $[\mathbf{G}_{kj}(\boldsymbol{\lambda}_{kj})]^{-1}$  é a inversa generalizada de ordem  $q_{kj} \times q_{kj}$  da matriz simétrica  $\mathbf{G}_{kj}(\boldsymbol{\lambda}_{kj})$ .

Se na Equação (2.2) não existe os efeitos aleatórios, logo resume-se o modelo à

$$\begin{aligned} \mathbf{Y} &\stackrel{ind}{\sim} D(\boldsymbol{\mu}, \boldsymbol{\sigma}, \boldsymbol{\nu}, \boldsymbol{\tau}) \\ \boldsymbol{\eta}_1 &= g_1(\boldsymbol{\mu}) = \mathbf{X}_1\boldsymbol{\beta}_1 \\ \boldsymbol{\eta}_2 &= g_2(\boldsymbol{\sigma}) = \mathbf{X}_2\boldsymbol{\beta}_2 \\ \boldsymbol{\eta}_3 &= g_3(\boldsymbol{\nu}) = \mathbf{X}_3\boldsymbol{\beta}_3 \\ \boldsymbol{\eta}_4 &= g_4(\boldsymbol{\tau}) = \mathbf{X}_4\boldsymbol{\beta}_4. \end{aligned} \tag{2.3}$$

Stasinopoulos et al. (2017) aponta o modelo da Equação (2.3) como o GAMLSS paramétrico e a Equação (2.2) como modelo GAMLSS com efeito aleatório. Quando temos um modelo GAMLSS paramétrico, ou seja, que não existem funções de suavizações, exige só a estimação de  $\beta$ . Já quando temos um modelo GAMLSS com efeitos aleatórios exige também  $\gamma$  e  $\lambda$  determinado por,

$$\boldsymbol{\lambda} = (\boldsymbol{\lambda}_{11}^T, \dots, \boldsymbol{\lambda}_{1J_1}^T, \boldsymbol{\lambda}_{21}^T, \dots, \boldsymbol{\lambda}_{4J_4}^T)^T.$$

Os GAMLSS paramétricos na Equação (2.3) pode ser analisados também na área dos modelos de delineamento experimentais (FISHER et al., 1937; MONTGOMERY, 2017), no qual modela-se o parâmetro de locação, porém com o ganho de poder modelar o parâmetro de escala, simetria e curtose. Os modelos GAMLSS paramétricos são ajustados pelas estimativas de máxima verossimilhança em relação à estimação de  $\beta$ . Enquanto o modelo GAMLSS não paramétrico com efeitos aleatórios é ajustado pelo método de estimação por máxima verossimilhança penalizada, em relação a estimação de  $\beta$  e  $\gamma$  para um  $\lambda$  fixo.

Para ajustar modelos GAMLSS, existe dois algoritmos básicos que são aplicados, tais como o algoritmo *backfitting*. O primeiro é proposto por Cole e Green (1992) que a partir de uma generalização tornou-se o algoritmo CG. Ele propõe o uso das primeiras derivadas e valores esperados ou aproximados das derivadas de segunda ordem e derivadas cruzadas da função de verossimilhança no que se refere aos parâmetros da distribuição  $\boldsymbol{\theta} = (\boldsymbol{\mu}, \boldsymbol{\sigma}, \boldsymbol{\nu}, \boldsymbol{\tau})$  com quatro parâmetros. Mas para muitas funções de densidade de probabilidade  $f(\boldsymbol{\gamma} | \boldsymbol{\theta})$ , os parâmetros  $\boldsymbol{\theta}$  dispõem de informação ortogonal, isto é, que os valores das derivadas cruzadas da função de verossimilhança são iguais à zero. Sendo assim, utilizamos um segundo algoritmo mais simples e que não utiliza o valor esperado das derivadas cruzadas, chamado RS, que é uma generalização do algoritmo proposto por Rigby e Stasinopoulos (2005). Existe ainda uma terceira metodologia, que une passos dos dois algoritmos, dando início por RS e finalizando com CG (ROQUIM, 2018).

Se tratando do algoritmo RS o seu objetivo é maximizar a verossimilhança em cada parâmetro  $(\boldsymbol{\mu}, \boldsymbol{\sigma}, \boldsymbol{\nu}, \boldsymbol{\tau})$  até atingir a convergência. E já o algoritmo CG, temos que a partir das derivadas cruzadas desde que não sejam nulas, irá atualizar os parâmetros em conjunto com cada iteração. Esta é a diferença fundamental entre eles, vale ressaltar que, em geral o

algoritmo RS é mais estável e rápido, quando se é comparado ao CG (STASINOPOULOS et al., 2017).

### 2.1.3.2 Inferência

Quando se trata de modelos paramétricos, a estimação dos parâmetros da Equação (2.3) é realizada via máxima verossimilhança, já se tratando em modelos GAMLSS semi-paramétricos da Equação (2.2), juntos a termos de suavização, é aplicada a máxima verossimilhança penalizada. Nos modelos paramétricos é importante que o modelo tenha um bom ajuste, para que via o método da máxima verossimilhança, se disponha assintoticamente, que todos estimadores dos parâmetros sejam consistentes, acompanhados de erros padrões assintóticos corretos, conduzido assim à cobertura correta dos intervalos de confiança. Porém, caso o modelo GAMLSS não seja adequado, isto é, desfavorável aos dados, logo os parâmetros podem ser inconsistentes, levando a inferências errôneas (STASINOPOULOS et al., 2017).

No modelo contém componentes não-paramétricos, por isso, importante conhecer a forma das curvas ajustadas para o termo de suavização e seus respectivos erros padrões. A estimação do parâmetros de suavização ( $\lambda$ ) é de total relevância porque indicará o formato de curva de suavização. Não é comum o interesse em erros padrões deste parâmetro, mas é possível que eles sejam aproximados por técnicas de *bootstrapping* (ROQUIM, 2018).

Stasinopoulos et al. (2017) atribuem os métodos inferenciais utilizados para resolver as questões citadas anteriormente sobre a classe da inferência baseada em verossimilhança e a classe em *bootstrapping*. A princípio um GAMLSS paramétrico e  $\boldsymbol{\theta}$  o vetor genérico de parâmetros. Diante disso, surge a ideia de um grupo com todos os coeficientes lineares para  $\mu, \sigma, \nu$  e  $\tau$ , por exemplo,  $(\beta_1, \beta_2, \beta_3, \beta_4)$ . A teoria de máxima verossimilhança clássica considerada, assintoticamente, é:

$$\hat{\boldsymbol{\theta}} \sim N(\boldsymbol{\theta}_T, \mathbf{i}(\boldsymbol{\theta}_T)^{-1}),$$

em que  $\hat{\boldsymbol{\theta}}$  é o estimador de máxima verossimilhança e

$$\mathbf{i}(\boldsymbol{\theta}_T) = -E \left[ \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^T} \right]_{\boldsymbol{\theta}_T},$$

sendo,  $(\mathbf{i}(\boldsymbol{\theta}_T))$  a matriz de informação esperada de Fisher tomando valores verdadeiros de  $\boldsymbol{\theta}_T$ . As vezes não é possível derivar a matriz de informação esperada analiticamente. Logo, utiliza a informação observada de Fisher, dada por

$$\mathbf{I}(\boldsymbol{\theta}_T) = - \left[ \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^T} \right]_{\boldsymbol{\theta}_T}.$$

Nota-se que  $\mathbf{I}(\boldsymbol{\theta}_T)$  é equivalente ao negativo da matriz hessiana do logaritmo da verossimilhança em  $\boldsymbol{\theta}_T$ . A matriz de variâncias e covariâncias da distribuição assintótica

de  $\hat{\theta}$  é aproximado por  $\mathbf{I}(\theta_T)^{-1}$  ao contrário de  $\mathbf{i}(\theta_T)^{-1}$ . O  $\theta_T$  substitui por  $\hat{\theta}$  quando  $\theta_T$  for desconhecido, recorrendo às matrizes de informações esperadas e observadas, dadas por  $\mathbf{i}(\hat{\theta})$  e  $\mathbf{I}(\hat{\theta})$ , respectivamente (STASINOPOULOS et al., 2017).

Sendo assim, para GAMLSS paramétricos, a distribuição assintótica para  $(\hat{\theta})$  é a seguinte,

$$\hat{\theta} \sim (\theta_T, \mathbf{I}(\hat{\theta})^{-1}).$$

Os erros padrões dos parâmetros estimados são obtidos em geral pela raiz quadrada da diagonal da matriz de variâncias e covariâncias associada à  $\theta$ . Para obter os erros padrões (EP) de  $\hat{\beta}$ , no caso de a matriz de informação para todos os parâmetros ser de difícil obtenção, recomenda-se a solução sugerida por Aitkin et al. (2009):

$$EP(\hat{\beta}) \approx \frac{|\hat{\beta}|}{\sqrt{\Delta GDEV}},$$

em que  $\Delta GDEV$  é a diferença no desvio global obtida pela omissão das variáveis explicativas associadas ao coeficiente do modelo o  $\beta$ . Este procedimento é justificado por aproximações da estatística de teste de razão de verossimilhanças à estatística do teste de Wald, sob a hipótese nula de que  $\beta = 0$ , em que (STASINOPOULOS et al., 2017)

$$\left[ \frac{\hat{\beta}}{EP(\hat{\beta})} \right]^2 \approx \Delta GDEV.$$

### 2.1.3.3 Tipos de distribuições dentro da família GAMLSS

Existem mais de 200 distribuições existentes na classe do GAMLSS e instaladas no pacote *gamlss* do R. Ao escolher qual a distribuição poderá ser utilizada depende da variável resposta que está sendo modelada. Dentro da família GAMLSS há três tipos distintos de famílias de distribuições:

- i. Distribuições contínuas;
- ii. Distribuições Discretas;
- iii. Mistura de Distribuições.

Para maiores informações, bem como uma tabela com os nomes das distribuições, espaço paramétrico e domínio das funções, respectivas funções de ligação e nome da distribuição na implementação do R, podem ser vistos em Stasinopoulos et al. (2017) páginas 148 - 151.

### 2.1.3.4 Critério de seleção dos modelos

As técnicas de seleção dos modelos de regressão tem o intuito de resolver o problema de selecionar preditores adequados dentre uma infinidade de possíveis preditores em potencial. A seleção de modelos GAMLSS inclui a melhor distribuição para a variável resposta, dos preditores adequados para os parâmetros da distribuição selecionada, das funções e dos hiperparâmetros.

O critério de Akaike generalizado (GAIC) e este leva em consideração o número de parâmetros e dos graus de liberdade utilizados no modelo para penalizar os modelos mais complexos e evitar ajustes aos dados em amostras de grandes tamanhos (PAIVA; FREIRE; CECATTI, 2008). Ele é utilizado no mesmo contexto do critério de informação de Akaike (AIC) (AKAIKE, 1998). O GAIC é definido por Voudouris et al. (2012), como:

$$GAIC(k) = -2l(\hat{\theta}) + (k \times gl),$$

em que  $l$  é o logaritmo da função verossimilhança e  $gl$  são os graus efetivos de liberdade do modelo ajustado,  $k$  é constante e  $gl$  é a penalidade para cada grau de liberdade utilizado. Refere-se a  $-2l(\hat{\theta})$ , como desvio global, pois GAIC ( $k$ ) é a estatística obtida pela adição do desvio global.

O desvio global (GDEV), importante medida para a seleção de modelos em GAMLSS, tal define-se como,

$$GDEV = -2l(\hat{\theta}),$$

em que,  $l(\hat{\theta})$  é o logaritmo da função de verossimilhança ajustada.

O modelo com o menor valor do critério GAIC ( $k$ ), para algum  $k$  escolhido é então selecionado e considerado melhor ajuste, pois GAIC ( $k$ ) penalizada modelos com muitos parâmetros. O critério (AIC) (AKAIKE, 1998) e o critério de informação bayesiano (BIC) (SCHWARZ, 1978) são casos especiais do GAIC ( $k$ ) um caso se  $k = 2$  recai no (AIC) e um outro caso  $k = \ln(n)$  recai no (BIC) (VOUDOURIS et al., 2012).

A seleção da distribuição adequada é feita em duas etapas sendo elas: o estágio de ajuste e o do diagnóstico. O de ajuste envolve comparação de diferentes modelos GAMLSS ajustados usando o GAIC e o de diagnóstico que é feito graficamente via análise dos gráficos de resíduos e dos *worm plot*.

### 2.1.3.5 Análise de resíduos

Os resíduos nos permitem extrair informações que diagnosticam ou demonstram a qualidade do ajuste de um modelo (etapa de diagnóstico), além de averiguar se as suposições do modelo foram satisfeitas. Muitas ferramentas gráficas são utilizadas para identificar incongruências entre o modelo ajustado e as observações coletadas, um equilíbrio

entre essas duas fontes de informação é necessário para que o modelo possa ser analisado com algum sentido prático.

A vantagem principal dos resíduos quantílicos (aleatorizados) normalizados é que, para qualquer distribuição da variável resposta, os resíduos verídicos sempre terão uma distribuição normal padrão quando o modelo assumido está correto. A averiguação de pressupostos por meio da normalidade dos resíduos já é bastante estabelecido na literatura estatística. Porém, os resíduos quantílicos (aleatorizados) normalizados fornecem-nos uma maneira comum de verificar a adequação de um modelo (STASINOPOULOS et al., 2017).

A análise utiliza os resíduos dos quantis aleatórios normalizados que podem ser utilizados para analisar exclusivamente a distribuição utilizada para ajustar tal modelo, introduzidos (DUNN P.K; SMYTH, 1996), é definido por,

$$r_i = \Phi^{-1} \{F(y_i; \hat{\theta})\},$$

em que  $\Phi^{-1}$  denota a inversa da função de distribuição acumulada de uma normal padrão,  $F(\cdot)$  é a função de distribuição acumulada adequada aos dados e  $\theta$  é o vetor de parâmetros. Observe que, um modelo adequado tem seus resíduos  $r_i$  seguindo a distribuição Normal Padrão, para o modelo ser considerado correto.

Uma das técnicas aplicadas para analisar os resíduos de um GAMLSS são os gráficos de minhoca (*worm plot*), o estágio de diagnóstico envolve o uso deste gráfico bem específico, que permite a detecção de inadequações no modelo globalmente ou dentro do intervalo de uma variável explicativa (BUUREN; FREDRIKS, 2001; BUUREN, 2007). O *worm plot* funciona como o gráfico normal *quantil-quantil* (Q-Q) sem tendência a fim de evidenciar alterações locais no domínio de uma dada covariável.

Os gráficos *worm plots* permitem identificar intervalos em que o modelo não é bem ajustados aos dados, seu eixo vertical descreve para cada observação, a diferença entre à sua localização nas distribuições teórica e empírica. Os pontos observados em conjunto, representam uma curva que se assemelha a uma minhoca. Esta forma gráfica aponta como os dados se distanciam da distribuição assumida e, quando tomadas em conjunto, sugere modificações úteis no modelo, como pode ser observado na Tabela 1 as interpretações de diferentes *worm plot*. Logo, se os pontos estão situados no interior da banda de confiança de 95% (Entre as duas curvas elípticas), o ajuste do modelo é considerado satisfatório.

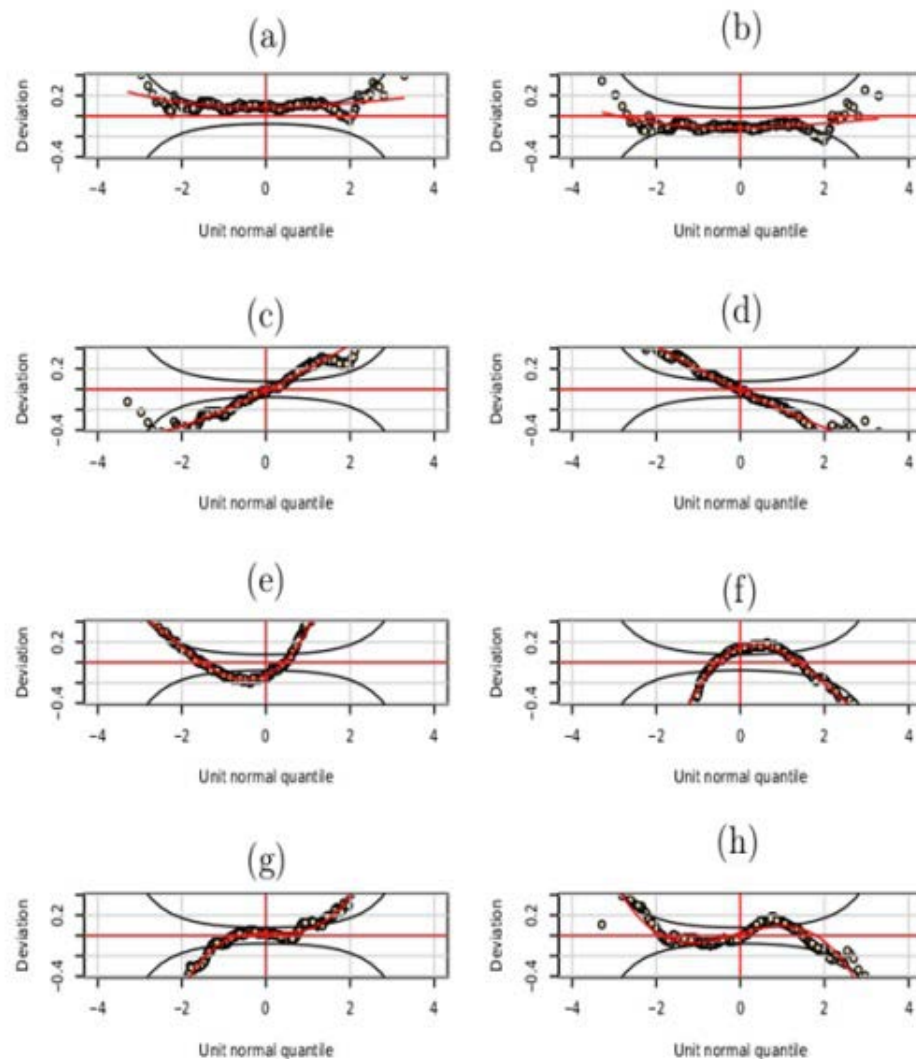


Tabela 1 – Interpretação de diferentes formatos de *worm plot*

Gráfico	Forma	Resíduos	Parâmetro ajustado
(a)	curva acima da origem	média muito alta	locação subestimada
(b)	curva abaixo da origem	média muito baixa	locação superestimada
(c)	curva inclinação positiva	variância muito alta	escala subestimada
(d)	curva inclinação negativa	variância muito baixa	escala superestimada
(e)	curva U	assimetria positiva	assimetria subestimada
(f)	curva U invertido	assimetria negativa	assimetria superestimada
(g)	S com curva esquerda para baixo	leptocurtose	calda muito leve
(h)	S com curva esquerda para cima	platicurtose	calda muito pesada

Fonte: Stasinopoulos et al. (2017).

Como pode ser observado na Tabela 1, as diferentes interpretações para os possíveis formatos dos pontos abrangendo desde variações na média (alta e baixa), variância (alta e baixa), assimetria (positiva e negativa) e curtose (caldas leves e pesadas). As interpretações auxiliam na tomada de decisão quanto ao ajuste do modelo aos dados, indicando se há necessidade de se escolher outro modelo probabilístico e/ou a inclusão de algum termo aos parâmetros de locação, escala e forma. Graficamente este comportamento pode ser observado na Figura 1.

Figura 1 – Diferentes problemas que podem ser diagnosticados pelo *worm plot*.

Fonte: Stasinopoulos et al. (2017).

Os diagnósticos por meio de gráficos na Figura 1 apontam como os dados se distanciam da distribuição assumida e, quando tomadas em conjunto, sugerem modificações úteis no modelo, a interpretação pode ser combinada com as informações da Tabela 1. O *worm plot* é uma ferramenta bastante útil para verificar adequabilidade e qualidade de ajuste. Desde que a modelagem via GAMLSS foi implementada este gráfico vem sendo utilizado por diversos autores (MACERA, 2011; ROQUIM, 2018).

### 3 APLICAÇÃO

A instituição Gapminder foi fundada em Estocolmo em 25 de fevereiro de 2005 por Ola Rosling, Anna Rosling Rönnlund e Hans Rosling. Ela é também uma empresa sem fins lucrativos com a intenção de promover desenvolvimento global. Oferece serviços como projetos colaborativos com universidades, organizações da ONU, agências públicas e organizações não governamentais (GAPMINDER, Acesso em: 29 de março 2022). É uma organização dedicada a educar o público através de dados para dissipar informações sobre o mundo em desenvolvimento. Ela usa dados para mostrar como as tendências atuais nos campos de saúde e da economia se contradiz com as informações soltas nas mídias sobre catástrofe, tragédias e outros eventos.

#### 3.1 Caracterização do estudo

Neste estudo foi considerado o banco de dados do pacote *gapminder* disponível no *software* estatístico R. Este banco contém dados históricos com 1704 observações e seis variáveis sendo elas, expectativa de vida, população total, PIB per capita (Produto interno bruto) dados em unidades de dólares internacionais, ano disposto em 1952-2007 incrementado de cinco em cinco anos, continente e país. Para o mesmo iremos ajustar a variável expectativa de vida com relação a PIB per capita. Portanto, considerou-se um modelo da classe dos modelos de regressão, sob a suposição que as observações seguem uma distribuição da classe dos modelos GAMLSS, com descrição da seguinte forma,

$$Y_{ij} = \beta_0 + \beta_1 x^2 + \sigma \varepsilon_{ij},$$

em que  $Y_{ij}$  é o valor observado em cada renda  $i$ ,  $\beta_0$  o efeito do intercepto e  $\beta_1$  o valor de cada  $i$  e  $\varepsilon_{ij} \sim BCPE(\mu, \sigma, \nu, \tau)$  representam os efeitos do fator não controlado do estudo, com  $i = 1, \dots, I$  e  $j = 1, \dots, J$ , tal que  $I$  denota o número de observações na variável e  $J$  o número de repetições.

#### 3.2 Distribuições de probabilidade

As distribuições consideradas foi na classe GAMLSS paramétrico (2.3) no qual é dada da seguinte forma,

$$\mathbf{Y} \stackrel{ind}{\sim} D(\boldsymbol{\mu}, \boldsymbol{\sigma}, \boldsymbol{\nu}, \boldsymbol{\tau})$$

$$\boldsymbol{\eta}_1 = g_1(\boldsymbol{\mu}) = \mathbf{X}_1 \boldsymbol{\beta}_1$$

$$\boldsymbol{\eta}_2 = g_2(\boldsymbol{\sigma}) = \mathbf{X}_2 \boldsymbol{\beta}_2$$

$$\boldsymbol{\eta}_3 = g_3(\boldsymbol{\nu}) = \mathbf{X}_3 \boldsymbol{\beta}_3$$

$$\boldsymbol{\eta}_4 = g_4(\boldsymbol{\tau}) = \mathbf{X}_4 \boldsymbol{\beta}_4.$$

A Distribuição Exponencial Potência de Box-Cox (BCPE), inserida por Rigby e Stasinopoulos (2004) para estimativa de percentil, trás um modelo para a variável resposta  $Y$  exibindo tanto assimetria quanto curtose. A distribuição é composta por quatro parâmetros  $(\mu, \sigma, \nu, \tau)$ , é um caso particular da Box-Cox Normal (RIGBY; STASINOPOULOS, 2005). Ela é especificada para a variável aleatória positiva  $Y$  através da variável aleatória transformada  $Z$  dada por,

$$Z = \begin{cases} \frac{1}{\sigma\nu} \left[ \left( \frac{Y}{\mu} \right)^\nu - 1 \right], & \text{se } \nu \neq 0. \\ \frac{1}{\sigma} \log \left( \frac{Y}{\mu} \right), & \text{se } \nu = 0, \end{cases} \quad (3.1)$$

para  $0 < Y < \infty$ , onde  $\mu > 0$ ,  $\sigma > 0$ ,  $-\infty < \nu < \infty$  e a variável aleatória  $Z$  é assumida por uma distribuição exponencial de potência padrão com parâmetro de potência,  $\tau > 0$ , tratado como parâmetro contínuo. A parametrização da Equação (3.1) foi usada por Cole e Green (1992) que assumiram uma distribuição normal padrão para  $Z$ .

A função de densidade de probabilidade de  $Z$ , é dada por:

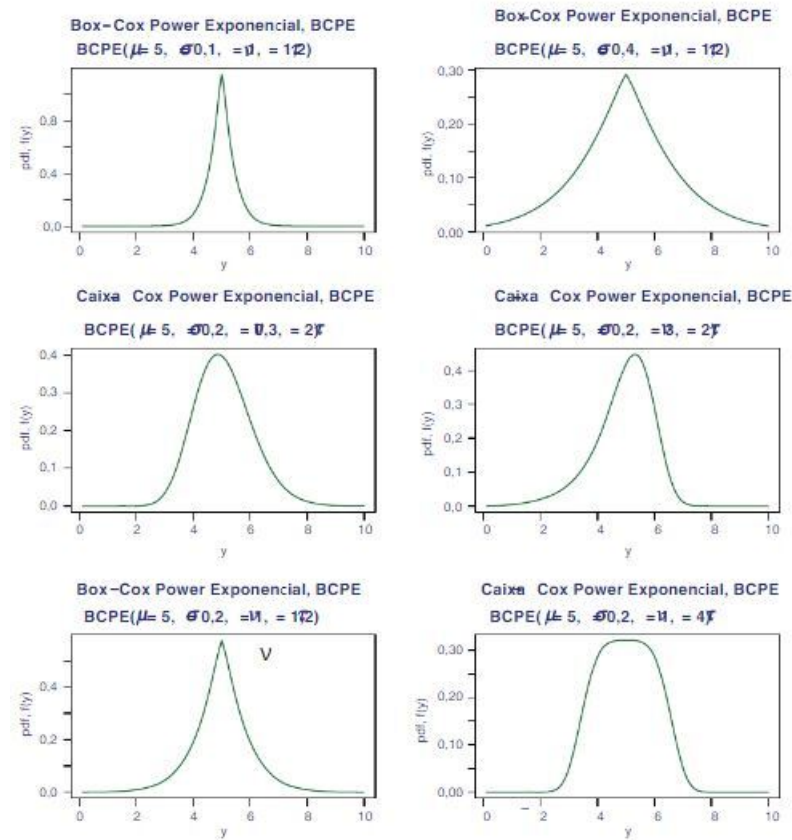
$$f_Z(z) = \frac{\tau}{c^{2(1+1/\tau)} \Gamma(1/\tau)} \exp \left\{ -0.5 \left| \frac{z}{c} \right|^\tau \right\}, \quad (3.2)$$

tal que  $-\infty < z < \infty$  e  $\tau > 0$ , onde  $c^2 = 2^{-2\tau} \Gamma(1/\tau) [\Gamma(3/\tau)]^{-1}$ . Essa parametrização garante que  $Z$  tenha média zero e desvio padrão 1 para todo  $\tau > 0$ . A distribuição exata da Equação (3.1) é uma distribuição exponencial de potência truncada, a partir dela a função de densidade de probabilidade de  $Y$  é dada por,

$$f_Y(y) = f_Z(z) \left| \frac{dz}{dy} \right| = \frac{y^{\nu-1}}{\mu^\nu \sigma} f_Z(z), \quad (3.3)$$

para  $y > 0$ . Os parâmetros  $\mu, \sigma, \nu, \tau$  podem ser interpretados como relativos à localização (mediana), escala (coeficiente de variação), assimetria (transformação para simetria) e curtose (parâmetro exponencial de potência), respectivamente. A Figura 2 mostra como seus valores afetam a curva da distribuição BCPE.

Figura 2 – Exemplos teóricos da Distribuição Exponencial Potência de Box-Cox



Fonte: Voudouris et al. (2012).

A Figura 2 mostra seis gráficos da distribuição BCPE para diferentes valores dos parâmetros. Todos os seis gráficos têm o mesmo  $\mu$ . Os dois primeiros gráficos mostram o efeito do  $\sigma$  (mantendo os outros três parâmetros fixos), ou seja, um aumento na propagação da distribuição; os dois gráficos do meio mostram o efeito de  $\nu$ , ou seja, mudança na assimetria, e os dois gráficos últimos gráficos mostram o efeito de  $\tau$ , ou seja, mudança na curtose.

Algumas outras distribuições foram ajustadas como a Normal (N), Normal acrescida do parâmetro  $\sigma$  (N +  $\sigma$ ), Gama Generalizada (GG) e Box Cox t (BCT). As análises foram feitas no *software* R (versão 4.1.0). Utilizou-se o pacote *gamlss* para as análises e a função *fitDist* foi utilizada para primeiramente escolher as distribuições probabilísticas que trouxesse melhores resultados. Aplicou-se o critério de GAIC com  $k = 1$ , para auxiliar na escolha do melhor modelo. Além dos gráficos de diagnósticos dos resíduos que permitem extrair as informações que diagnosticam ou demonstra a qualidade do ajuste do modelo e verificar se as suposições do modelo foram satisfeitas utilizou em conjunto o gráfico *worn plot* que admite observar incongruências entre o modelo ajustado e os dados.

E para um modelo ser considerado adequado seus resíduos devem seguir a distribuição Normal padrão  $N(0, \sigma^2)$ , sendo assim também foi considerado o teste de Shapiro-Wilks

com a seguinte estatística de teste,

$$W = \frac{b^2}{\sum_{i=1}^n (x_i - \bar{x})^2},$$

com  $x_i$  o valor de cada medida independente,  $\bar{x}$  a média de todas as medidas e  $b$  é um parâmetro que deve ser calculado a partir da quantidade de observações ( $n$ ). Considerando as seguintes hipóteses que são testadas com um nível de significância ( $\alpha(\%)$ ) para um determinado p-valor,

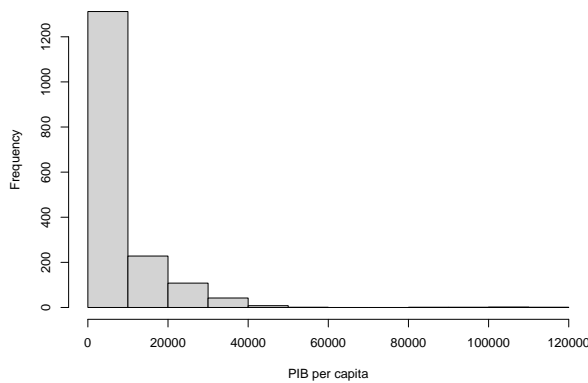
$$\begin{cases} H_0 : & \text{Os resíduos tem distribuição normal} \\ H_1 : & \text{Os resíduos não tem distribuição normal,} \end{cases}$$

temos que, se o p-valor  $\leq \alpha$ , rejeita-se  $H_0$ , ou seja, rejeita-se a hipótese de que os resíduos seguem distribuição normal (SHAPIRO; WILK, 1965).

## 4 RESULTADOS E DISCUSSÕES

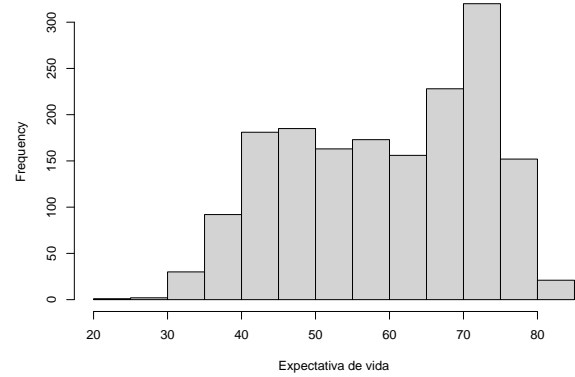
Os resultados desenvolvidos serão apresentados na sequência mostrando as técnicas adotadas. Primeiramente foi ajustado um histograma aos dados de expectativa de vida e PIB per capita, o que se pode ser observado nas Figura 3 e Figura 4, respectivamente. Há evidências que os dados não são normalmente distribuídos, indicando que dificilmente uma distribuição normal, por exemplo, irá ajustar os dados adequadamente.

Figura 3 – Histograma do PIB per capita



Fonte: Elaborado pelo Autor.

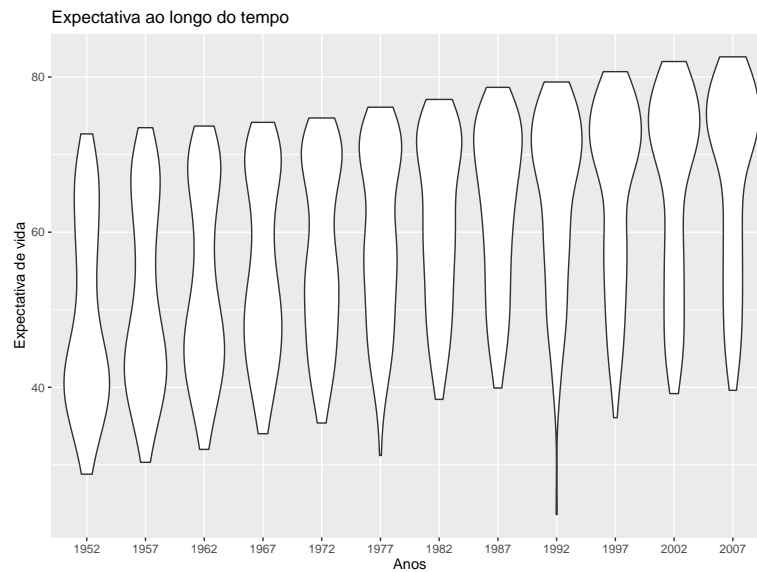
Figura 4 – Histograma da expectativa de vida



Fonte: Elaborado pelo Autor.

O gráfico de dispersão no formato violino apresentado na Figura 5 trás uma útil visualização para o comportamento da expectativa de vida, pois é possível enxergar a variabilidade e assimetria dentro de cada ano. Sendo assim, tem-se no eixo das abscissas os anos que estão distribuídos de cinco em cinco anos, e no eixo das ordenadas as expectativas de vida. Ao observamos é possível dizer que a medida que os anos se passam as caudas superiores se alargam, representando maior densidade de pontos, ou seja, existe uma maior frequência de dados. O gráfico mostra que a partir do ano 1977 é possível visualizar que a expectativa de vida aumentou em comparação aos anos anteriores.

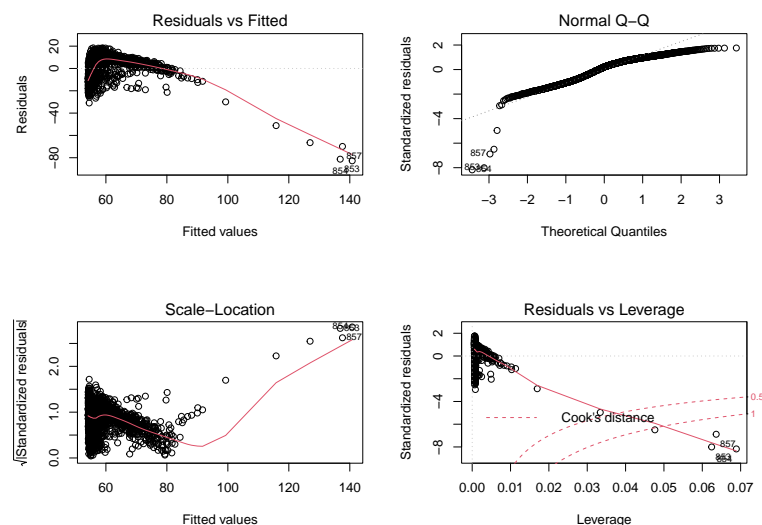
Figura 5 – Gráfico de dispersão da expectativa de vida ao longo do tempo



Fonte: Elaborado pelo Autor.

O ajuste do modelo de regressão linear sob suposição de distribuição normal para os resíduos são apresentados na Figura 6. De acordo com o gráfico quantil quantil percebe-se que os resíduos não seguem uma distribuição Normal, mostrando que a modelagem padrão não se adapta aos dados, fato que corrobora com a Figura 4. Além disso, foi aplicado o teste de Shapiro-Wilks com um P valor igual a  $2,2e - 16$ , ou seja, menor que 0,05 o que indica a ausência de normalidade dos resíduos , ao nível de 5% de significância.

Figura 6 – Gráfico de resíduos para Expectativa de Vida



Fonte: Elaborado pelo Autor.

Após todas as técnicas realizadas e não obter resultados satisfatórios, optou-se



por utilizar a técnica de Modelos Aditivos Generalizados para Localização, Escala e Forma (GAMLSS) (RIGBY; STASINOPOULOS, 2005), já que esta apresenta flexibilidade na modelagem dos dados, sendo uma forte alternativa para ajuste dos dados com grandes dificuldades. Na Tabela 2, foi ajustado distribuições candidatas à modelar os dados via comando *fitDist* dentro do pacote *gamlss*, considerando os reais positivos.

Tabela 2 – Principais distribuições dos modelos probabilísticos ajustados via pacote *gamlss* utilizando o Critério de Akaike Generalizados

Distribuição	GAIC	Distribuição	GAIC
BCPE <sub>o</sub>	13170,71	WEI3	13483,58
BCPE	13170,1	exGAUS	13560,29
GG	13384,58	NO	13558,29
GB2	13386,58	GA	13633,04
BCCG <sub>o</sub>	13438,65	GIG	13635,04
BCCG	13438,65	IG	13696,85
BCT <sub>o</sub>	13440,65	LOGNO2	13699,62
BCT	13440,65	LOGNO	13699,62
WEI2	13483,58	IGAMMA	13783,39
WEI	13483,58	EXP	17333,54
GP	17335,54	PARETO2 <sub>o</sub>	17335,54

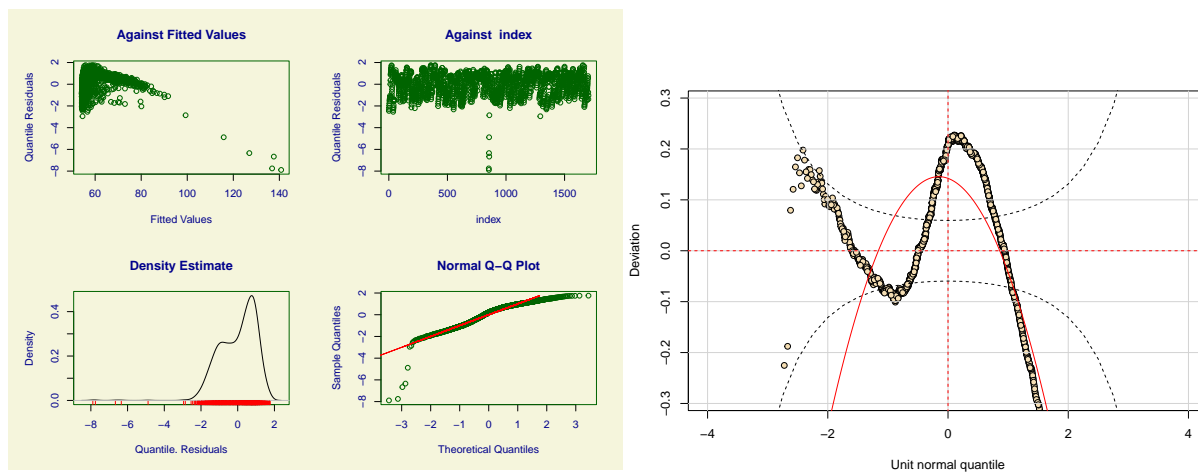
Fonte: Elaborado pelo Autor.

Dentre as distribuições obtidas e apresentadas na Tabela 2, observa-se que a distribuição de PARETO2<sub>o</sub> não foi uma das melhores candidatas para modelar os dados e que as distribuições Normal, Gama Generalizada e Box Cox t se encaixam em intermediárias levando em consideração os valores de GAIC ordenados do menor para o maior. Enquanto a distribuição BCPE foi a que menor teve valor de GAIC, sendo a melhor das distribuições candidatas. Sendo assim, ao ajustar os dados foi utilizadas as distribuições: Normal, a Normal acrescida do parâmetro de escala ( $\sigma$ ) (Normal +  $\sigma$ ), Gama Generalizada, Box Cox t e Exponencial Potência de Box-Cox. Para esses modelos ao tentar ajustar os parâmetros de forma( $\nu, \tau$ ) não convergia, mas ainda assim foi possível encontrar um modelo que adequasse satisfatoriamente aos dados.

Inicialmente modelou-se a distribuição Normal, mas os dados não ajustaram satisfatoriamente como observado na Figura 7. Na tentativa de prosseguir com a Normal foi acrescentado o parâmetro ( $\sigma$ ), mas não se adequou aos dados como pode ser visto na Figura 8. Para efeitos de comparações foram aplicadas as distribuições GG (Figura 9) e a BCT (Figura 10), porém, observando os gráficos os dados não foram ajustados satisfatoriamente. Portanto, com o resultado do comando *fitDist* que teve o melhor valor de GAIC, aplicou-se a distribuição Exponencial Potência de Box-Cox (BCPE) considerando um tratamento polinomial grau 2 no parâmetro de localização, e que ao observar os gráficos da Figura 11 obteve um dos melhores ajuste para os dados e, também de acordo com

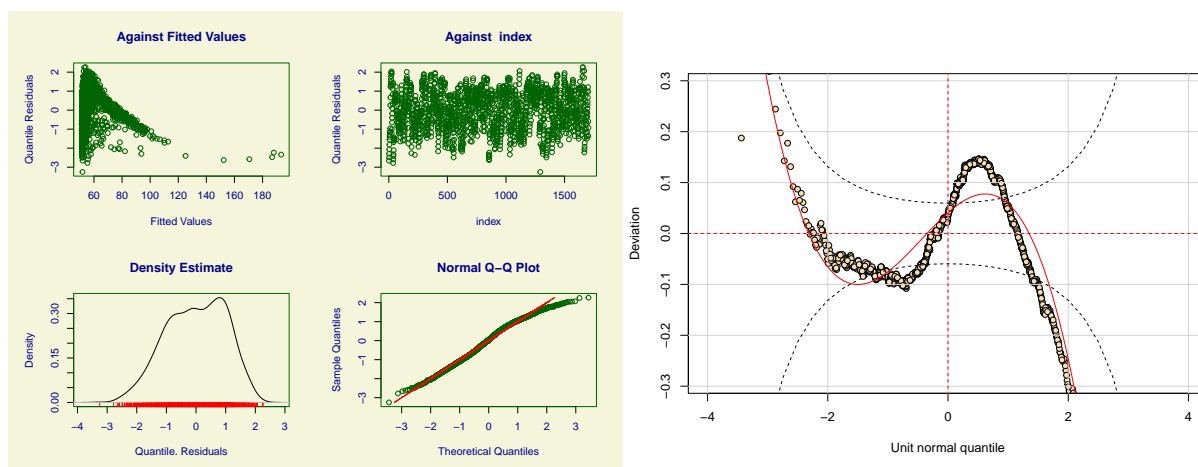
o critério de GAIC obteve o melhor/menor resultado em relação as outras distribuições ajustadas.

Figura 7 – Gráficos residuais e *Worm plot* da Distribuição Normal



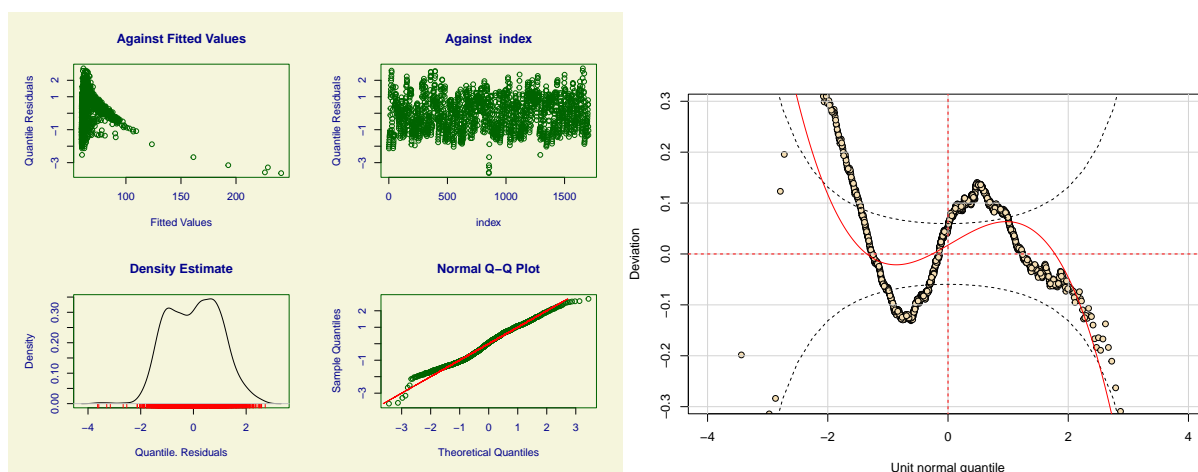
Fonte: Elaborado pelo Autor.

Figura 8 – Gráficos residuais e *Worm plot* da Distribuição Normal +  $\sigma$



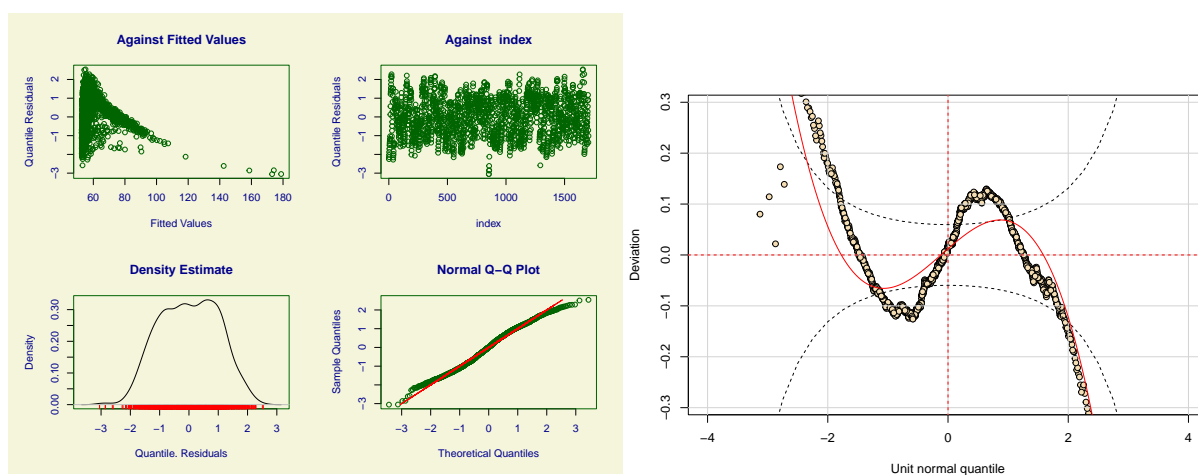
Fonte: Elaborado pelo Autor.

Figura 9 – Gráficos residuais e *Worm plot* da Distribuição Gama Generalizada



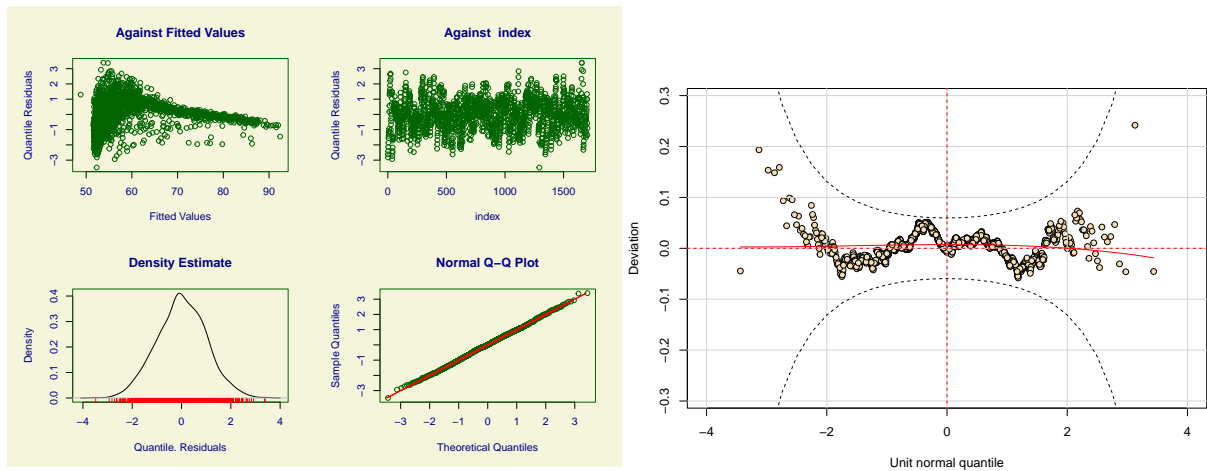
Fonte: Elaborado pelo Autor.

Figura 10 – Gráficos residuais e *Worm plot* da Distribuição Box Cox t



Fonte: Elaborado pelo Autor.

Figura 11 – Gráficos residuais e *Worm plot* da Distribuição Exponencial Potência de Box-Cox



Fonte: Elaborado pelo Autor.

Os resultados apresentados nos gráficos de resíduos e *Worm plots* na Figura 11 do modelo (M5) evidenciam os melhores resultados do que os encontrados nos outros quatro modelos. Existe um ponto no modelo (M5) que viola as bandas de confiança, porém este não corresponde a 5% dos pontos que poderiam estar fora do intervalo com 95% de confiança.

Na Tabela 3 está descrito as distribuições com as diferentes configurações paramétricas e os valores de GAIC para cada uma ajustada.

Tabela 3 – Descrição dos principais modelos ajustados

Distribuições	GAIC	$\mu$	$\sigma$	$\nu$	$\tau$
M1: Normal (N)	12850,41	X	-	-	-
M2: Normal (N)+ $\sigma$	12585,83	X	X	-	-
M3: Gama Generalizada (GG)	12652,91	X	X	-	-
M4: Box Cox T (BCT)	12607,46	X	X	-	-
M5: Exponencial Potência de Box-Cox (BCPE)	12345,19	X*	-	-	-

X indica o parâmetro considerado/X\* indica tratamento no parâmetro

Fonte: Elaborado pelo Autor.

Na Tabela 4 tem-se as estatísticas obtidas para o modelo (M5) com média 0,006012 bem próxima a zero, a variância 0,999083 bem próxima a 1 logo está aproximadamente de uma normal padrão, assimetria igual a  $-0,007958$  relativamente próxima do zero, curtose 2,979493 aproximadamente igual a 3 e o Coeficiente de curtose de Filliben igual a 0,999674 próxima de 1. Por fim, foi aplicado o teste de Shapiro Wilks nos resíduos do modelo (M5), obtendo um P Valor igual a 0,83 que indica presença de normalidade dos dados ao nível de

5% de significância. Com esses resultados trás mais segurança para o ajuste na aplicação da distribuição dos dados.

Tabela 4 – Estimativas dos parâmetros para o ajuste da Distribuição Exponencial Potência de Box-Cox

Parâmetro	Valores estimados
Média	0,006012
Variância	0,999083
Coefficiente de assimetria	-0,007958
Coefficiente de curtose	2,979473
Coefficiente de curtose de Filliben	0,999674

Fonte: Elaborado pelo Autor.

Ao analisar as estimativas da Tabela 4 conjuntamente com os gráficos da Figura 11, pode-se observar que a Distribuição Exponencial Potência de Box-Cox foi a que melhor se adequou aos dados dentre as que foram ajustadas. Uma das contribuições deste trabalho é que para as curvas de crescimento e modelagens de forma geral, o GAMLSS é uma metodologia bastante flexível que vem sendo pesquisada e utilizada com frequência. Curvas de referências ligadas ao comportamento do crescimento de fetos, para valores de parâmetros antropométricos na população foram utilizados por Paiva, Freire e Cecatti (2008), os resultados encontrados indicam que a classe de modelos GAMLSS é alusão para o estudo de curvas de referências.

Kienen e Taconeli (2015) estudaram uma classe de modelos para ajustar as curvas de crescimento dos custos de procedimentos hospitalares de uma operadora de planos de saúde em um estudo realizado no município de Curitiba-PR em 2013. Os fatores como o tipo de parto, local de destino e tipo de acomodação mostraram-se associados aos custo dos procedimentos. O médico responsável pelo parto foi incluído aos modelos por meio de recursos aleatórios, permitindo identificar a contribuição dos médicos para o custo final dos partos e avaliar tudo entre os efeitos. Uma das metodologias adotadas foi utilizando uma abordagem baseada na Distribuição Exponencial Potência de Box Cox (BCPE) e as comparações com este modelo foram feitas assumindo menores valores de AIC e BIC e também com a normalidade dos resíduos. Os resultados revelaram que o modelo com base na BCPE obtiveram um dos melhores resultados de acordos com os critérios determinados como o critério de Akaike, tendo sido escolhido, de forma que os autores usaram critérios parecidos com os utilizados neste trabalho.

Portanto, a classe de modelos de regressão GAMLSS permite que distintos tipos de efeitos aleatórios sejam inclusos na formação dos modelos, sendo assim foi realizado um estudo com experimentos de dados de crescimento de plantas usando a metodologia GAMLSS com efeito aleatório e comparado a uma abordagem de modelo linear misto, com resultados melhores para o modelo GAMLSS sob a Distribuição Gama Generalizada (GG) (THOMAS et al., 2018).

## 5 CONCLUSÃO

Modelos de regressão nem sempre atendem as condições necessárias para ajustar o comportamento dos dados, o que mostra a importância da aplicação da metodologia GAMLSS na modelagem. Neste trabalho foram apresentadas às vertentes relacionadas a inferência e diagnóstico trazendo a flexibilidade que as análises do modelo permitem por possuir uma ampla família de distribuições para a variável resposta. Ajustar os parâmetros de média e dispersão, além de incluir os suavizadores, retorna resíduos bem comportados, o que evidencia um ajuste satisfatório.

Portanto, o uso dos Modelos Aditivos Generalizados para Localização, Escala e Forma (GAMLSS) foi adequado para modelar os dados do pacote *gapminder* utilizando as variáveis expectativa de vida em relação ao PIB per capita, representando a versatilidade desta metodologia nas mais diversas áreas de aplicação da estatística. Os modelos da classe GAMLSS trouxe os melhores resultados que as distribuição Normal, com pelo menos cinco modelos de ajustes. No qual o modelo (M5) aplicado a Distribuição Exponencial Potência de Box-Cox (BCPE), que nos oferece a possibilidade de ajustar a curtose e assimetria dos dados, foi o melhor considerando os parâmetros modelados com efeitos da variável independente. Por fim, a principal ideia desse estudo foi ter conseguido gerar as curvas de crescimento por meio do método GAMLSS que permitiu flexibilidade ao formato da modelagem.

## REFERÊNCIAS

- AITKIN, M. et al. *Statistical modelling in R*. [S.l.]: Oxford University Press Oxford, 2009. Citado na página 20.
- AKAIKE, H. Information theory and an extension of the maximum likelihood principle. In: *Selected papers of hirotugu akaike*. [S.l.]: Springer, 1998. p. 199–213. Citado na página 21.
- BUUREN, S. v. Worm plot to diagnose fit in quantile regression. *Statistical Modelling*, Sage Publications India Pvt. Ltd, B-42, Panchsheel Enclave, New Delhi, v. 7, n. 4, p. 363–376, 2007. Citado na página 22.
- BUUREN, S. v.; FREDRIKS, M. Worm plot: a simple diagnostic device for modelling growth reference curves. *Statistics in medicine*, Wiley Online Library, v. 20, n. 8, p. 1259–1277, 2001. Citado na página 22.
- CASELLA G.; BEERGER, R. L. . inferência estatística. 2 Ed. São Paulo: Cengage Learning, 588 p, 2010. Citado na página 13.
- COLE, T. J.; GREEN, P. J. Smoothing reference centile curves: the lms method and penalized likelihood. *Statistics in medicine*, Wiley Online Library, v. 11, n. 10, p. 1305–1319, 1992. Citado 2 vezes nas páginas 18 e 26.
- DEMÉTRIO, C. G. B. *Modelos lineares generalizados em experimentação agrônômica*. [S.l.]: USP/ESALQ, 2001. Citado na página 15.
- DRAPER N. E SMITH, H. *Applied Regression Analysis*. [S.l.]: 3 Ed. New York, USA: John Wiley e Sons, 736 p, 1998. Citado na página 13.
- DUNN P.K; SMYTH, G. Resíduos de quantis aleatórios. *Jornal de Estatísticas Computacionais e Gráficas*, v. 5, 1996. Citado na página 22.
- EILERS P. H.; MARX, B. D. D. M. Twenty years of p-splines. *SORT: statistics and operations research transactions*, v. 39, 2015. Citado na página 15.
- FISHER, R. A. et al. The design of experiments. *The design of experiments*., Oliver & Boyd, Edinburgh & London., n. 2nd Ed, 1937. Citado na página 18.
- GANZOLA, S. Construção de um modelo de regressão para avaliação de imóveis. *Dissertação (Mestrado em Engenharia de Produção) – Universidade Federal de Santa Catarina*, Florianópolis, SC, 2002. Citado na página 13.
- GAPMINDER. Disponível em: <https://www.gapminder.org/about/about-gapminder/history/>. Acesso em: 29 de março 2022. Citado na página 25.
- HASTIE T. E TIBSHIRANI, R. Generalized additive models. [S.l.]: Wiley Online Library, 1990. Citado 3 vezes nas páginas 11, 12 e 15.
- HOFFMANN, R. *Análise de Regressão: Uma introdução à econometria*. [S.l.]: Hucitec, 2006. Citado na página 13.
- KIENEN, R.; TACONELI, C. Modelos generalizados aditivos para locação, escala e forma numa análise de custos de procedimentos hospitalares de uma operadora de planos de saúde. v. 33, p. 330–342, 09 2015. Citado na página 35.

- MACERA, M. A. C. *Uso dos métodos clássico e bayesiano para os modelos não-lineares heterocedásticos simétricos*. Tese (Doutorado) — Universidade de São Paulo, 2011. Citado na página 24.
- MONTGOMERY. *Design e análise de experimentos*. [S.l.]: John wiley & filhos, 2017. Citado na página 18.
- NELDER J. A. E WEDDERBURN, R. W. M. Generalized linear models. *Journal of the Royal Statistical Society, v. 135, n. 3, p. 370–384*, v. 135, 1972. Citado 3 vezes nas páginas 11, 12 e 14.
- PAIVA, C. S. M.; FREIRE, D. M. C.; CECATTI, J. G. Modelos aditivos generalizados para posição, escala e forma (gamlss) na modelagem de curvas de referência. *Rev. bras. ciênc. saúde*, p. 289–310, 2008. Citado 2 vezes nas páginas 21 e 35.
- PNUD, O. Programa das nações unidas para o desenvolvimento. *Atlas do desenvolvimento humano no Brasil*, 2012. Citado na página 11.
- RENCHER A. C.; SCHAALJE, G. B. *Linear models in statistics*. [S.l.: s.n.], 2008. Citado na página 14.
- RIGBY, R. A.; STASINOPOULOS, D. M. Smooth centile curves for skew and kurtotic data modelled using the box–cox power exponential distribution. *Statistics in medicine, Wiley Online Library*, v. 23, n. 19, p. 3053–3076, 2004. Citado na página 26.
- RIGBY, R. A.; STASINOPOULOS, D. M. Generalized additive models for location, scale and shape. *Journal of the Royal Statistical Society: Series C (Applied Statistics), Wiley Online Library*, v. 54, 2005. Citado 5 vezes nas páginas 12, 16, 18, 26 e 31.
- ROQUIM, F. V. Gamlss na experimentação agropecuária: um estudo em doenças parasitárias de bovinos de leite. *Dissertação (Estatística e Experimentação Agropecuária) – Universidade Federal de Lavras. Lavras*, 2018. Citado 3 vezes nas páginas 18, 19 e 24.
- SCHWARZ, G. Estimating the dimension of a model. *The annals of statistics*, JSTOR, p. 461–464, 1978. Citado na página 21.
- SEERS, D. O desafio às teorias e estratégias desenvolvimentistas. *Revista Brasileira de Economia*, v. 24, n. 3, p. 5–28, 1970. Citado na página 11.
- SHAPIRO, S. S.; WILK, M. B. An analysis of variance test for normality (complete samples)†. *Biometrika*, v. 52, n. 3-4, p. 591–611, 12 1965. ISSN 0006-3444. Disponível em: <<https://doi.org/10.1093/biomet/52.3-4.591>>. Citado na página 28.
- STASINOPOULOS, M. D. et al. *Flexible regression and smoothing: using GAMLSS in R*. [S.l.]: CRC Press, 2017. Citado 8 vezes nas páginas 15, 17, 18, 19, 20, 22, 23 e 24.
- TEAM, R. C. R: A language and environment for statistical computing. *Vienna, Austria, Disponível em: <http://www.R-project.org/>*, 2018. Citado na página 16.
- THOMAS, G. et al. Analysis of a longitudinal multilevel experiment using gamlss. *arXiv preprint arXiv:1810.03085*, 2018. Citado na página 35.
- VOUDOURIS, V. et al. Modelling skewness and kurtosis with the bcpe density in gamlss. *Journal of Applied Statistics*, Taylor & Francis, v. 39, n. 6, p. 1279–1293, 2012. Citado 2 vezes nas páginas 21 e 27.