

ESTIMATIVAS DA POPULAÇÃO DE CRUZ DAS ALMAS-BAHIA, EM 2020, UTILIZANDO OS AJUSTES LINEAR, EXPONENCIAL E SIGMOIDAL

do Nascimento, P. H. R. *

Latrilha, J. V. L. †

Revista Eletrônica de Ciências Exatas e Tecnológicas

MAR-2022, 3ª Edição, Volume 1.

Submitted: 22 dez.2021. Approved: 03 fev.2022.

RESUMO

Neste trabalho, apresentamos um pequeno estudo sobre os modelos de Malthus e de Verhulst, precursores do estudo da dinâmica populacional. Utilizamos o Método dos Mínimos Quadrados, com o objetivo de obter os ajustes por uma reta de regressão e por duas curvas: a exponencial e a sigmoide, dos dados disponibilizados pelo IBGE da população da cidade de Cruz das Almas. Além disso, vista a ausência de realização do Censo IBGE, no ano de 2020, foi feita uma comparação da estimativa dada pelo IBGE com os ajustes obtidos.

Palavras-chave: Dinâmica Populacional. Regressão Linear. Ajuste Exponencial. Ajuste Sigmoide. Cruz das Almas.

ABSTRACT

In this work, we present a small study on the Malthus and Verhulst models, precursors of the study of population dynamics. We used the Least Squares Method, in order to obtain the adjustments by a regression line and by two curves: the exponential and the sigmoid, from the data provided by the IBGE for the population of the city of Cruz das Almas. In addition, given the lack of completion of the IBGE Census in 2020, a comparison was made between the estimate given by the IBGE with the adjustments obtained.

Keywords: Population Dynamics. Linear Regression. Exponential Adjustment. Sigmoidal Adjustment. Cruz das Almas.

Sumário

Sumário	1
Introdução	2
FUNDAMENTAÇÃO TEÓRICA	3
Equações Diferenciais	3
Modelos Populacionais	4
Modelo de Malthus	4
Modelo de Verhulst	5
Ajuste de Curvas	7
Ajuste por uma reta	7
Ajuste por uma curva exponencial	8
Ajuste por uma curva sigmoide	8
APLICAÇÃO: AJUSTES E ESTIMATIVAS	8
Ajuste Linear	8
Análise do Ajuste Linear	9
Gráfico dos Valores Ajustados x Resíduos	9
Gráfico dos Quantis Teóricos x Resíduos	9
Gráfico Escala-Localização	10
Gráfico Alavancagem x Resíduos	10
Ajuste Exponencial	11
Ajuste Sigmoide	11
CONSIDERAÇÕES FINAIS	12

*Paulo Henrique Ribeiro do Nascimento. Licenciado em Matemática pela Universidade Católica do Salvador (UCSal) e Mestre em Matemática pela Universidade Federal da Bahia (UFRB). Filiação: Centro de Ciências Exatas e Tecnológicas (CETEC)/ Universidade Federal do Recôncavo da Bahia (UFRB), Brasil. E-mail: nascimento.p@ufrb.edu.br

†João Victor Luz Latrilha. Graduando no curso de Bacharelado em Ciências Exatas e Tecnológicas pela Universidade Federal do Recôncavo da Bahia (UFRB), Brasil. E-mail: victorlatrilha@gmail.com

Referências	12
APÊNDICE A Códigos Python	14

INTRODUÇÃO

Por causa da importância do estudo da dinâmica de populações para o desenvolvimento de uma sociedade, há na literatura uma quantidade significativa de trabalhos que abordam o assunto e, por isso, mencionaremos alguns trabalhos.

Bezerra et al. (2016) realizou uma projeção do número de habitantes da cidade de Bataguassu - MS segundo o modelo de Malthus Discreto e Contínuo, comparando os valores estatísticos do Instituto Brasileiro de Geografia e Estatística (IBGE) do período de 1980 à 2010 e sua projeção até o ano de 2020. Pôde-se notar que os resultados são próximos, porém apresentam variação considerável com relação aos dados obtidos pelo IBGE, isso se deve ao fato do modelo de Malthus considerar a taxa de crescimento populacional *per capita* constante e ilimitado. Logo, concluiu-se que o modelo se torna mais eficiente quando aplicado a um curto período de tempo.

Dias e Pessoa (2016) trataram, matematicamente, da evolução do crescimento populacional da cidade de São José dos Campos - SP. Foram utilizados os dados populacionais da cidade disponibilizados no site da Universidade Estadual Paulista (UNESP) e do IBGE, no período de 1920 a 2015. Foi feita a análise dos dados populacionais por meio dos modelos clássicos de Malthus e Verhulst e chegou ao resultado que ambos os modelos representaram o crescimento exponencial de forma próxima aos dados disponibilizados no período de 1920 a 1996. Porém, a partir do ano de 1996, a população passou a apresentar um crescimento linear demonstrando uma taxa de crescimento fixa em aproximadamente 10.000 habitantes por ano.

Tavares (2017), que realizou um estudo sobre as tendências de crescimento e decréscimo da população de trabalhadores do Polo Naval e da quantidade de habitantes do Município de Rio Grande - RS, entre os anos de 2005 e 2015. Para avaliar a dinâmica da população da região, foi utilizado o ajuste por Mínimos Quadrados junto aos

modelos analíticos clássicos de Malthus e Verhulst, utilizando o software wxMaxima. Após os estudos, concluiu-se que os modelos mostram, em geral, resultados próximos da realidade e podem auxiliar na análise da evolução do crescimento dessas populações.

Neste trabalho, faremos uma estimativa da população da cidade de Cruz das Almas, em 2020, vista a não realização do Censo até a data atual. Para isso, apresentaremos os modelos de Malthus e de Verhulst e ajustaremos os dados do IBGE através da reta de regressão e das curvas exponencial e sigmoide.

O município de Cruz das Almas foi criado através da Lei nº 119 de 29 de julho de 1897, desmembrando-se de São Félix. Os primeiros habitantes do município vieram em boa parte de cidades vizinhas como Cachoeira e São Félix, motivados pela riqueza do seu solo. A origem do seu nome faz menção a uma lenda antiga onde os tropeiros da época passavam pela região que ao chegarem à antiga vila de Nossa Senhora do Bonsucesso, se deparavam com uma cruz em frente a igreja matriz no centro da vila. Neste mesmo local, eles paravam e rezavam para as almas de seus mortos (HERBÁRIO, 2021).

O Censo é uma pesquisa que reúne os dados estatísticos dos habitantes de uma cidade, província, estado ou nação. O Censo no Brasil tem sua pesquisa realizada através de visitas em todos os domicílios (por volta de 58 milhões) espalhados pelo território nacional ($8.514.876,599\text{km}^2$) e tem como objetivo, conhecer a situação de vida da população. Se trata de um trabalho gigantesco que envolve cerca de 230 mil pessoas. Tem por objetivo dar subsídios para a definição de políticas públicas em nível nacional, estadual e municipal e também ajuda a iniciativa privada. Além disso, é possível acompanhar o crescimento, a distribuição geográfica e a evolução de outras características da população ao longo do tempo (IBGE, 2010). Já Cunha (2021) nos diz que o Censo obtém os dados necessários para realizar a distribuição dos recursos públicos para cada cidade, além planejar políticas públicas para solucionar possíveis problemas onde for necessário.

No Brasil, é considerado pelo IBGE que o primeiro Censo foi realizado no ano 1872, época em que o país era uma monarquia governada pelo imperador D. Pedro II. Além de perguntas sobre sexo, idade, escolaridade e estado civil, religião e deficiência física, o questionário incluía a pergunta se a pessoa era livre ou escrava. Já o primeiro Censo a ser realizado pelo IBGE foi o de 1940. A periodicidade do censo demográfico no Brasil se deu através da lei nº 8.184, de 10 de maio de 1991 (BRASIL, 1991). Entre os especialistas é inquestionável a importância de se respeitar o intervalo de dez anos, pois a recomposição de uma população nacional com base em Censos posteriores ou anteriores não é tarefa fácil, nem em países com população estável. Além disso, trata-se de uma operação dispendiosa e que necessita de grande preparo logístico.

Segundo Weterman (2021), estava prevista a realização de um novo Censo Demográfico em 2020 mas, por conta da pandemia, este foi adiado para o ano de 2021. Entretanto, após a Comissão Mista de Planos, Orçamentos Públicos e Fiscalização ¹ aprovar a redução em 96% dos recursos destinados para a realização da pesquisa, sua execução ficou inviável.

FUNDAMENTAÇÃO TEÓRICA

Nessa seção, apresentaremos definições e resultados que envolvem as equações diferenciais ordinárias (EDO) (BOYCE W. E.; DIPRIMA, 2010) que serão utilizadas em dois modelos de dinâmica de populações, os modelos de Malthus e de Verhulst. Além desses dois modelos, apresentaremos uma breve revisão teórica da regressão linear com o intuito de fazer uma estimativa da população da cidade de Cruz das Almas, Bahia, no ano de 2020.

Equações Diferenciais

Uma equação que contém derivadas de uma (ou mais de uma) variável dependente com relação a apenas a uma variável independente é denominada **Diferencial Ordinária** (EDO). Caso a variável (ou variáveis) dependente(s) esteja(m) vin-

culada(s) a mais de uma variável independente, dizemos que a equação é **Diferencial Parcial** (EDP). A **ordem** de uma EDO é a ordem da maior derivada que aparece na equação. Podemos escrever uma EDO de ordem n com apenas uma variável dependente da seguinte maneira:

$$\underbrace{a_n \frac{dy^n}{dt^n} + a_{n-1} \frac{dy^{n-1}}{dt^{n-1}} + \dots + a_1 \frac{dy}{dt} + a_0 y}_{\mathcal{L}(y)} = g(t), \quad (1)$$

em que $a_k = a_k(t)$, $k = 0, \dots, n$ e $g(t)$ são funções que dependem apenas de $t \in I \subset \mathbb{R}$, onde I é um intervalo. Claramente, se $a_n(t) \neq 0, \forall t \in I$, podemos reescrever a equação da seguinte forma:

$$\frac{dy^n}{dt^n} + p_{n-1} \frac{dy^{n-1}}{dt^{n-1}} + \dots + p_1 \frac{dy}{dt} + p_0 y = G(t), \quad (2)$$

em que $p_k = p_k(t) = \frac{a_k(t)}{a_n(t)}$, $k = 0, \dots, n-1$ e $G(t) = \frac{g(t)}{a_n(t)}$.

Dizemos que uma EDO é **homogênea** se $g(t) = 0, \forall t \in I$, ou seja,

$$a_n \frac{dy^n}{dt^n} + \dots + a_1 \frac{dy}{dt} + a_0 y = 0. \quad (3)$$

Caso contrário, é dita **não homogênea**.

A EDO (1) é **linear** se o operador \mathcal{L} é linear na variável y , onde $\mathcal{L} : \mathcal{F} \rightarrow \mathcal{F}$, em que \mathcal{F} é o espaço das funções n vezes diferenciável com derivadas contínuas até a n -ésima ordem, definidas em I , caso contrário, dizemos que ela é uma equação diferencial **não-linear**.

Uma solução de (1) é qualquer função $y = f(t)$, definida no intervalo $I \subset \mathbb{R}$, que quando substituída nesta EDO, fica reduzida a uma equação identidade. Entretanto, nem sempre temos a solução de uma EDO de forma **explícita**, existem aquelas que se apresentam de forma **implícita**. Por exemplo, para $-2 < t < 2$, a relação $t^2 + y^2 - 4 = 0$ é uma solução implícita, pois ao derivá-la, obtemos $y' = \frac{-t}{y}$.

Se as funções $y_1, y_2, \dots, y_k, k \in \mathbb{N}$, todas definidas em um intervalo I , são soluções EDO linear homogênea, a combinação linear

$$y = c_1 y_1(t) + c_2 y_2(t) + \dots + c_k y_k(t), \quad (4)$$

¹ A Comissão Mista de Planos, Orçamentos Públicos e Fiscalização (CMO) é um órgão legislativo do Congresso Nacional e é composta por Deputados e Senadores. Foi instituída pelo art. 166, §1º, da Constituição Federal de 1988.

em que $c_1, c_2, \dots, c_k, k \in \mathbb{N}$ são constantes arbitrárias, é também solução da EDO linear, no intervalo I .

Dizemos que as funções y_1, y_2, \dots, y_k são **linearmente dependentes** (LD), em um intervalo I , se existe um conjunto de constantes c_1, c_2, \dots, c_k , nem todas nulas, tais que

$$c_1 y_1 + c_2 y_2 + \dots + c_k y_k = 0, \quad (5)$$

para todo $t \in I$. Caso contrário, são ditas **linearmente independentes** (LI).

A combinação linear da [Equação 4](#), representa o conjunto de todas as soluções de uma EDO linear homogênea (também chamada de solução geral) se $k = n$ e as funções y_1, y_2, \dots, y_n são LI. O resultado a seguir é utilizado para verificar se um conjunto de funções é LI.

Teorema 1. [(BOYCE W. E.; DIPRIMA, 2010, p. 172)] *O conjunto de funções $\mathcal{Y} = \{y_1, y_2, \dots, y_n\}$, definidas em um intervalo aberto I e deriváveis até a ordem $n - 1$ é linearmente independente se, e somente se,*

$$\begin{aligned} W(\mathcal{Y}) &= W(y_1, y_2, \dots, y_n) \\ &= \begin{vmatrix} y_1 & y_2 & \dots & y_n \\ y_1' & y_2' & \dots & y_n' \\ \vdots & \vdots & \ddots & \vdots \\ y_1^{(n-1)} & y_2^{(n-1)} & \dots & y_n^{(n-1)} \end{vmatrix} \\ &\neq 0, \forall t \in I. \end{aligned}$$

O determinante $W(\mathcal{Y})$ é chamado **Wronskiano**².

Um Problema de Valor Inicial (PVI) ou de Cauchy é um conjunto de equações formado por uma EDO de ordem n e $n - 1$ condições complementares que determinam em um mesmo valor atribuído à variável independente, o valor da função dependente e de suas derivadas até a ordem $n - 1$. Claramente, a solução para um PVI é uma função y que satisfaz a EDO e todas as $n - 1$ condições iniciais. A existência e unicidade de solução para um PVI pode ser encontrada em (SOTOMAYOR T.; MANUEL, 1979). Os PVI apresentados neste trabalho possuem uma EDO de primeira ordem e sua

solução geral é a combinação linear de apenas uma solução.

Modelos Populacionais

A dinâmica de populações trata das variações, no tempo e no espaço, das densidades e tamanhos de população. Seu estudo visa à melhor compreensão da variação do número de indivíduos de uma determinada população e também, dos fatores que a influenciam em tais variações. Para isso, é necessário o conhecimento das taxas em que se verificam perdas e ganhos de indivíduos e identificar os processos que regulam a variação da população (NUNES; TAKAHASHI, 2009).

Segundo Meira et al. (2014), numa busca de compreender determinados fenômenos da natureza e esclarecer aspectos sociais relacionados a sua sobrevivência, o homem estuda a dinâmica das populações através de equações que se aproximem de sua descrição. Historicamente, estas questões foram abordadas, sobretudo, na área das ciências humanas, com enfoque em estudos geográficos. Contudo, a partir do século XVIII, Thomas Malthus desenvolveu um modelo matemático para estudar o crescimento populacional mundial, cuja solução provém de uma equação ordinária simples. A partir daí, foi possível desenvolver estimativas relevantes para o contexto histórico em que a teoria foi criada. Outras teorias matemáticas surgiram através do modelo de Malthus, porém ele continua sendo uma das alternativas mais eficientes para o estudo de grupos e populações pequenas sem limitantes naturais.

Modelo de Malthus

Consideremos uma população definida por uma função P que relaciona o número de indivíduos $P(t)$, no instante t , e analisemos o problema de como varia esta população com o passar do tempo. Para podermos equacionar esse problema, consideraremos que, para um pequeno intervalo de tempo Δt houve uma variação $\Delta P(t) = P(t + \Delta t) - P(t)$ na quantidade de indivíduos. Obviamente, se ao final do intervalo de tempo a população cresceu, ΔP será positivo. Caso contrário, negativo. Os fatores comuns que causam variação no tamanho da popu-

² Homenagem ao matemático polonês Josef Wronski (1776-1853).

lação são as quantidades \mathfrak{B} de **nascimento**, \mathfrak{J} de **imigração**, \mathfrak{D} de **morte** e \mathfrak{E} de **emigração**, ou seja,

$$\Delta P = \mathfrak{B} + \mathfrak{J} - \mathfrak{D} - \mathfrak{E}. \quad (6)$$

Consideremos, agora, que a população P está isolada ($\mathfrak{J} = \mathfrak{E} = 0$). Esta hipótese é feita com o objetivo de minimizar o problema e chegarmos ao modelo pretendido. Portanto, para um intervalo de tempo Δt , temos:

$$\frac{\Delta P}{\Delta t} = \frac{\mathfrak{B} - \mathfrak{D}}{\Delta t}, \quad (7)$$

Vamos considerar ainda que existem, inicialmente, P indivíduos na população. Então, as taxas médias de nascimento e de morte, por unidade de tempo e por indivíduo da população, são dadas, respectivamente, pela multiplicação entre o inverso do número inicial de indivíduos da população e a razão entre o número de nascimentos por variação de tempo e; pela multiplicação entre o inverso do número inicial de indivíduos da população e pela razão entre o número de mortes por variação de tempo (chamadas, respectivamente, taxas de nascimento e de morte *per capita*), ou seja:

$$b = \frac{1}{P} \frac{\mathfrak{B}}{\Delta t} \quad \text{e} \quad \mathfrak{d} = \frac{1}{P} \frac{\mathfrak{D}}{\Delta t}.$$

Note, pelas suas próprias definições, que b é uma constante positiva e que a constante positiva \mathfrak{d} é menor ou igual a 1.

Tomando $\tau = b - \mathfrak{d}$, a equação (7) pode ser reescrita como:

$$\frac{1}{P} \frac{\Delta P}{\Delta t} = \tau \quad (8)$$

representando, assim, uma taxa *per capita* de crescimento da população se $\tau > 0$, e de decrescimento, se $\tau < 0$. Se $\tau = 0$, não há variação no número de indivíduos na população, ou seja, não nascem, nem morrem indivíduos; ou ainda, o número de indivíduos que nascem é igual ao número de indivíduos que morrem.

Em intervalos de tempos iguais, podemos considerar Δt como sendo a unidade relativa ao problema, a população inicial como sendo $P = P_0$, e a equação (8) reescrita como:

$$P = P_0(1 + \tau), \quad (9)$$

constituindo, assim, o modelo discreto de Malthus.

A sua formulação diferencial é obtida quando fazemos $\Delta t \rightarrow 0$ em (8), ou seja,

$$\lim_{\Delta t \rightarrow 0} \frac{1}{P} \frac{\Delta P}{\Delta t} = \lim_{\Delta t \rightarrow 0} \tau \implies \frac{1}{P} \frac{dP}{dt} = \tau. \quad (10)$$

Ao separarmos as variáveis e integrar a equação obtida, obtemos como solução de (10) a função

$$P(t) = \alpha e^{\tau t}, \quad (11)$$

em que α é uma constante de integração. Ao utilizarmos a condição $P(0) = P_0$ (população inicial), encontramos $\alpha = P_0$ e a solução da EDO (10) se torna

$$P(t) = P_0 e^{\tau t} \quad (12)$$

e, assim, um crescimento populacional exponencial para o modelo ($\tau > 0$) ou um decrescimento ($\tau < 0$).

Modelo de Verhulst

Como o modelo Malthusiano prevê um crescimento populacional exponencial, este modelo não serve para descrever dinâmicas populacionais para longos intervalos de tempo. Isso se deve ao assumir que a taxa de variação *per capita* de crescimento/decaimento é constante, independentemente do tamanho. Entretanto, o que na realidade é esperado é que, quanto maior fique uma população, a sua taxa de mortalidade cresça e a de natalidade, reduza. Tendo a taxa finita de crescimento de uma população

$$\tau = b - \mathfrak{d},$$

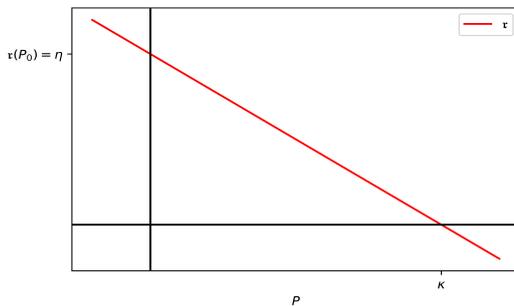
de maneira que se b diminui e \mathfrak{d} cresce, o valor de τ é uma função decrescente.

Claramente, existem infinitas funções decrescentes que descrevem esse comportamento. Mas vamos considerar a mais elementar destas, ou seja, tomar τ como uma função afim que depende do número de elementos da população, de modo que essa dependência seja de tal forma que a medida que P cresce, o valor de τ decresce, ou seja, tomemos a taxa *per capita* $\frac{1}{P} \frac{dP}{dt}$ de uma população para pequenos valores de P , como sendo igual a uma função afim τ onde, à medida que o valor de P cresce, $\tau(P)$ diminui (uma função afim decrescente).

Na dinâmica de populações, isso se justifica da seguinte maneira: para pequenas populações, os recursos no meio são abundantes e, à medida que a população cresce, em um determinado momento, sente-se a necessidade que essa taxa de crescimento diminua ou se anule.

A Figura 1 ilustra a função afim que modela a taxa de crescimento *per capita* de um modelo de dinâmica populacional que decresce à medida que a população aumenta.

Figura 1 – Taxa de Crescimento *per capita* afim



Fonte – Elaborada pelos autores

A equação da reta na Figura 1 é

$$\tau(P) = -\frac{\eta}{\kappa}P + \eta. \quad (13)$$

A partir das considerações feitas, temos que a equação do modelo se torna

$$\frac{dP}{dt} = \eta P \left(1 - \frac{P}{\kappa}\right), \quad (14)$$

chamado **modelo logístico**, uma equação não linear (no segundo membro aparece a variável P^2).

Os parâmetros κ e η possuem um interpretação para a dinâmica de uma população. Se observamos que, para valores de P menores que κ temos $\frac{dP}{dt} > 0$, ou seja, a população cresce. Por outro lado, se, para valores de P maiores que κ , $\frac{dP}{dt} < 0$, o que significa que a população diminui. Assim, o valor κ , que está associado ao valor $\tau(P) = 0$, é um valor crítico para o tamanho P da população, chamado de **capacidade de carga** do modelo, ou seja, representa o número máximo de indivíduos que o modelo pode suportar antes que a população comece a diminuir. Observamos, ainda, que para populações muito menores que κ , o fator $(1 - P/\kappa)$

da equação do modelo logístico se aproxima de 1 e, conseqüentemente, este modelo se aproxima do modelo Malthusiano e η pode ser interpretado como sendo a maneira como a população varia quando esta está muito abaixo da capacidade de carga.

A fim de determinar a solução da Equação 14, considere P uma população que segue o modelo:

$$\frac{1}{P} \frac{dP}{dt} = \eta(1 - \gamma P), \quad (15)$$

em que $\gamma = \kappa^{-1}$.

Efetuada a distribuição do fator η no segundo membro da Equação 15, dividindo-a por P e isolando, no segundo membro, o termo independente de P , obtemos:

$$-P^{-2} \frac{dP}{dt} + \eta P^{-1} = \eta \gamma. \quad (16)$$

Façamos $\varphi = P^{-1}$, o que implica em $\frac{d\varphi}{dt} = -P^{-2} \frac{dP}{dt}$ e, substituindo o que foi obtido na Equação 16, temos:

$$\frac{d\varphi}{dt} + \eta \varphi = \eta \gamma. \quad (17)$$

Ao multiplicar a Equação 17 por $e^{\eta t}$, observamos que o primeiro membro da equação obtida é o resultado da derivada do produto desta função pela função φ , ou seja,

$$\frac{d}{dt} [e^{\eta t} \varphi(t)] = e^{\eta t} \eta \gamma. \quad (18)$$

Integrando a Equação 18 e, em seguida, substituindo φ por P^{-1} , temos a solução geral da Equação 14:

$$P(t) = \frac{1}{\gamma + C_1 e^{-\eta t}}. \quad (19)$$

Tomando $P(0) = P_0$, temos que

$$P(0) = \frac{1}{\gamma + C_1 e^{-\eta 0}} \Rightarrow C_1 = \frac{1}{P_0} - \gamma$$

Assim,

$$P(t) = \frac{1}{\gamma + \left(\frac{1}{P_0} - \gamma\right) e^{-\eta t}}. \quad (20)$$

Observamos que se $t \rightarrow \infty$, então $P(t) \rightarrow \frac{1}{\gamma} = \kappa$. Além disso, se $P_0 < \kappa$, $\frac{dP}{dt} = \eta P(1 - \gamma P) >$

0. Sendo assim,

$$\begin{aligned} \frac{d^2 P}{dt^2} &= \frac{d}{dt}(\eta P - \gamma \eta P^2) \\ &= \eta \frac{dP}{dt} - 2\gamma \eta P \frac{dP}{dt} \\ &= \eta(1 - 2\gamma P) \frac{dP}{dt}. \end{aligned}$$

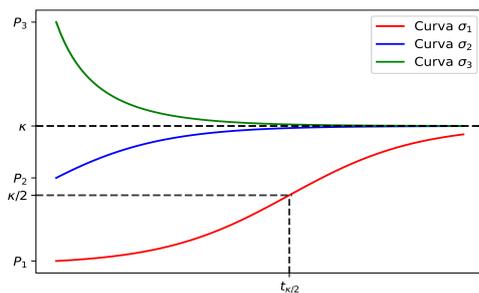
Como $\frac{dP}{dt} > 0$ e $\eta > 0$, o estudo do sinal de $\frac{d^2 P}{dt^2}$ depende apenas de $(1 - 2\gamma P)$. Logo,

$$\frac{d^2 P}{dt^2} = 0 \Leftrightarrow (1 - 2\gamma P) = 0 \Leftrightarrow P = \frac{1}{2\gamma} = \frac{\kappa}{2},$$

a ordenada do ponto de inflexão do gráfico de P .

A seguir, ilustramos através da Figura 2, diferentes situações dos gráficos de P para $0 < P_0 < \frac{\kappa}{2}$, $\frac{\kappa}{2} < P_0 < \kappa$ e $P_0 > \kappa$.

Figura 2 – Soluções do Modelo de Verhulst



Fonte – Elaborada pelos autores: Código Python 9

Utilizando a Equação 20, a Figura 2 mostra que, para:

- $P_0 = P_1$, com $0 < P_1 < \frac{\kappa}{2}$ (Curva σ_1), existe um crescimento mais rápido da população até o tempo $t_{\kappa/2} = \frac{-1}{\eta} \ln\left(\frac{\kappa}{P_1} - 1\right)$ e, depois deste, o crescimento se torna mais lento.
- $P_0 = P_2$, com $\frac{\kappa}{2} < P_2 < \kappa$ (Curva σ_2), a população apresenta, apenas, crescimento mais lento.
- $P_0 = P_3$, com $P_3 > \kappa$ (Curva σ_3), o número de indivíduos da população, para se ajustar à capacidade de suporte do meio (razão entre

o que é produzido e a taxa de consumo *per capita*), deve diminuir.

Em qualquer um dos casos, temos que o número de indivíduos da população tende à capacidade de carga do meio.

Ajuste de Curvas

Ajuste por uma reta

Quando verificamos que uma distribuição de dados, no plano, possui uma tendência linear, utilizamos o Método dos Mínimos Quadrados (MMQ)³ para obter a equação da reta que melhor se ajusta a esse conjunto de pontos (HOFFMAN; KUNZE, 1971), determinada pelos valores dos coeficientes angular e linear da reta.

Considere a distribuição

$$R = \{(x_k, y_k) \in \mathbb{R}^2, 1 \leq k \leq n\}$$

e sejam

$$X = \begin{bmatrix} x_1 & x_2 & \dots & x_n \end{bmatrix}^t$$

e

$$Y = \begin{bmatrix} y_1 & y_2 & \dots & y_n \end{bmatrix}^t$$

as representações matriciais, de ordem $n \times 1$ das abscissas e ordenadas de R . Vamos obter os valores das ordenadas Z dos pontos de uma reta que estão próximos dos valores de Y da distribuição, utilizando o MMQ. Então, definindo

$$\begin{aligned} Z &= \alpha_1 X + \alpha_2 \mathcal{J} \\ &= \begin{bmatrix} x_1 \alpha_1 + \alpha_2 \\ x_2 \alpha_1 + \alpha_2 \\ \vdots \\ x_n \alpha_1 + \alpha_2 \end{bmatrix} = \begin{bmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{bmatrix} \underbrace{\begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix}}_{\alpha}, \end{aligned}$$

em que $\mathcal{J} = \begin{bmatrix} 1 & 1 & \dots & 1 \end{bmatrix}^t$ é uma matriz de ordem $n \times 1$, \bar{X} é de ordem $n \times 2$ e α de ordem 2×1 .

Dessa forma, vamos minimizar a função

$$F(\alpha) = (Y - Z)^2 = (Y - \bar{X}\alpha)^2$$

e, para tal, devemos determinar α que satisfaz a:

$$\begin{bmatrix} 0 \\ 0 \end{bmatrix} = \frac{\partial F}{\partial \alpha} = \frac{\partial Z}{\partial \alpha} \frac{\partial F}{\partial Z}.$$

³ técnica de otimização matemática onde uma de suas aplicações é encontrar a curva que melhor se ajusta a um conjunto de dados que minimiza a soma dos desvios quadráticos entre o valor estimado (este presente na curva) e os dados observados.

Uma vez que

$$\frac{\partial F}{\partial Z} = \frac{\partial}{\partial Z}(Y - Z)^2 = -2 \underbrace{(Y - Z)}_{\text{matriz } n \times 1} \text{ e}$$

$$\frac{\partial Z}{\partial \alpha} = \frac{\partial}{\partial \alpha}(\bar{X}\alpha) = \underbrace{\bar{X}^t}_{\text{matriz } 2 \times n},$$

segue que

$$\begin{bmatrix} 0 \\ 0 \end{bmatrix} = \frac{\partial F}{\partial \alpha} = -2\bar{X}^t(Y - Z) = -2\bar{X}^t(Y - \bar{X}\alpha),$$

ou seja,

$$\bar{X}^t Y = \bar{X}^t \bar{X} \alpha \implies \alpha = (\bar{X}^t \bar{X})^{-1} \bar{X}^t Y. \quad (21)$$

Claramente, se $x_i \neq x_j$, para algum $i \neq j$, então $\bar{X}^t \bar{X}$ é não-singular e os valores de α_1 e α_2 , após alguns cálculos, são dados por:

$$\alpha_1 = \frac{\sum_{k=1}^n x_k \sum_{k=1}^n y_k - n \sum_{k=1}^n x_k y_k}{\left(\sum_{k=1}^n x_k\right)^2 - n \sum_{k=1}^n x_k^2} \text{ e} \quad (22)$$

$$\alpha_2 = \frac{\sum_{k=1}^n y_k - \alpha_1 \sum_{k=1}^n x_k}{n}. \quad (23)$$

A equação da reta de regressão linear é, portanto, dada por:

$$y = \alpha_1 x + \alpha_2. \quad (24)$$

Ajuste por uma curva exponencial

Podemos ajustar um conjunto de pontos (x_k, y_k) a uma curva exponencial

$$y = \alpha e^{\alpha_1 x}. \quad (25)$$

considerando:

$$\ln(y) = \ln(\alpha e^{\alpha_1 x}) = \ln(\alpha) + \alpha_1 x.$$

Tomando $Y = \ln(y)$ e $\alpha_2 = \ln(\alpha)$, reduzimos o problema de ajustar o conjunto de pontos (x_k, y_k) a uma curva exponencial, ao de ajustar o conjunto de pontos (x_k, Y_k) a uma reta.

Ajuste por uma curva sigmoide

Vimos que o modelo de Verhulst (ver [Equação 20](#)) é dado por:

$$P(t) = \frac{1}{\frac{1}{\kappa} + \left[\frac{1}{P_0} - \frac{1}{\kappa}\right] e^{-\eta t}}, \quad \kappa = \frac{1}{\gamma}. \quad (26)$$

A linearização ocorre da seguinte maneira:

$$\frac{1}{P(t)} = \frac{1}{\kappa} + \left[\frac{1}{P_0} - \frac{1}{\kappa}\right] e^{-\eta t}$$

$$\frac{1}{P(t)} - \frac{1}{\kappa} = \left[\frac{1}{P_0} - \frac{1}{\kappa}\right] e^{-\eta t}$$

$$e^{\eta t} = \left[\frac{\kappa - P_0}{P_0 \kappa}\right] \left[\frac{\kappa P(t)}{\kappa - P(t)}\right]$$

$$\ln \left[\frac{P(t)}{\kappa - P(t)}\right] = \eta t + \ln \left[\frac{P_0}{\kappa - P_0}\right]$$

Tomando $Y = \ln \left[\frac{P(t)}{\kappa - P(t)}\right]$ e $\alpha_2 = \ln \left[\frac{P_0}{\kappa - P_0}\right]$, determinamos $\alpha_1 = \eta$ e $P_0 = \frac{\kappa e^{\alpha_2}}{1 + e^{\alpha_2}}$.

APLICAÇÃO: AJUSTES E ESTIMATIVAS

Faremos dois ajustes aos dados do IBGE ([Tabela 1](#)) através de uma reta e de uma curva exponencial. Em seguida, duas estimativas da população de Cruz das Almas, em 2020, utilizando estes ajustes e algumas análises.

Tabela 1 – População de Cruz das Almas (1940 - 2010)

ANO	HABITANTES
1940	19.336
1950	20.822
1960	23.941
1970	29.355
1980	37.352
1991	45.874
2000	53.055
2010	58.606

Fonte – (IBGE, 1940, p. 405), (IBGE, 1952, p. 15), (IBGE, 1960, p. 80), (IBGE, 1973, p. 770), (IBGE, 1983, p. 222), (IBGE, 1991, p. 48), (SDT; MBA, 2015, p. 2) e (IBGE, 2021)

Ajuste Linear

Com o intuito de obter a reta de regressão, inicialmente construímos as matrizes x_k e y_k com os respectivos valores das colunas Ano e População da [Tabela 1](#) no Python. Para isso, foi necessário utilizar a biblioteca NumPy e os comandos do [algoritmo 1](#). Os valores dos coeficientes da reta de

regressão r foram determinados utilizando o [algoritmo 2](#). Assim, a equação da reta de ajuste linear é, aproximadamente:

$$y = 604,86x - 1.158.631,17. \quad (27)$$

Os valores populacionais obtidos através da [Equação 27](#) são apresentados na [Tabela 2](#) e foram determinados utilizando o [algoritmo 3](#).

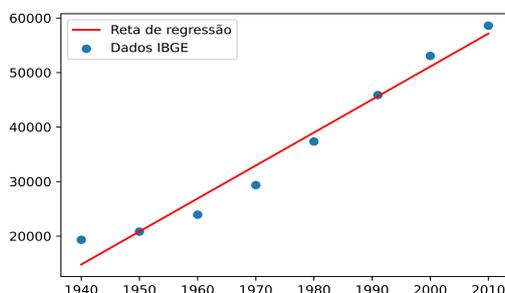
Tabela 2 – População de Cruz das Almas (1940 - 2010) segundo a reta de regressão

ANO	HABITANTES
1940	14.796
1950	20.845
1960	26.894
1970	32.942
1980	38.991
1991	45.644
2000	51.088
2010	57.137

Fonte – Elaborado pelos autores

Com o intuito de comparar os dados nas tabelas [Tabela 1](#) e [Tabela 2](#), vamos definir o desvio quadrático médio DQM como sendo a média dos desvios entre os valores reais e os estimados pela reta de regressão r . O desvio padrão DP é, aproximadamente, $DP = \sqrt{DQM} = 2.523$ habitantes. Este valor foi obtido utilizando o [algoritmo 4](#).

Figura 3 – População de Cruz das Almas (1940-2010) e a reta de regressão



Fonte – Elaborada pelos autores: Código Python 7

Além disso, através da reta de regressão r , podemos estimar que a população de Cruz das Almas é, em 2020, de 63.185 habitantes. Se com-

paramos este valor com o dado divulgado (IBGE, 2021), de 63.591 habitantes, temos um erro absoluto de 406 habitantes (aproximadamente 0,64%).

A [Figura 3](#) ilustra a os dados disponibilizados pelo IBGE e a reta de regressão obtida pela distribuição da [Tabela 1](#).

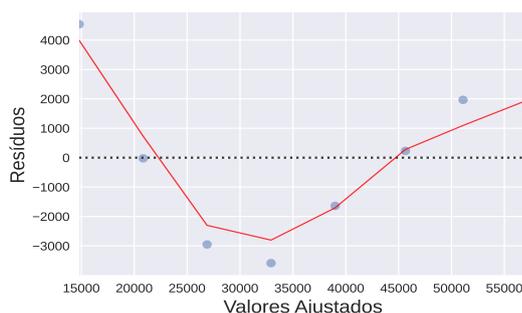
Análise do Ajuste Linear

Com o intuito de avaliar se o ajuste linear é bem qualificado, apresentaremos quatro gráficos de diagnóstico. Esses gráficos examinam diferentes suposições sobre a relação existente entre o modelo e os dados reais, a saber: se os dados podem ser ajustados por uma reta; se os erros são normalmente distribuídos, com média zero; se os erros têm variância constante (homocedasticidade) e; se não existem pontos de alta alavancagem.

Gráfico dos Valores Ajustados x Resíduos

Se os dados puderem ser ajustados por um modelo linear, então devemos ter uma linha relativamente horizontal ao observar o gráfico dos Valores Ajustados x Resíduos. A [Figura 4](#) mostra que não existe um padrão linear para os resíduos quando efetuamos um ajuste linear aos dados da [Tabela 1](#).

Figura 4 – Valores Ajustados x Resíduos



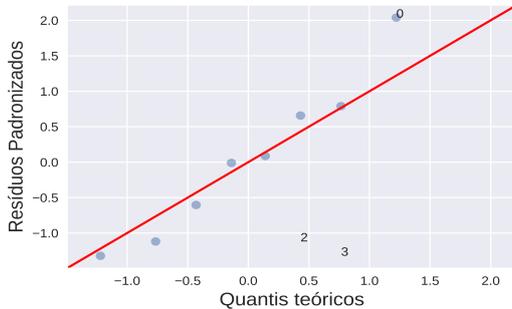
Fonte – Elaborada pelos autores: Código Python 11

Gráfico dos Quantis Teóricos x Resíduos

Este gráfico, também conhecido como Normal Q-Q, mostra se os resíduos estão normalmente distribuídos. Para isso, quanto mais próximos os pontos estiverem da primeira bissetriz, os resíduos estão sendo distribuídos normalmente e, portanto, é mais provável que vejamos valores esperados do que valores extremos. A [Figura 5](#) mostra que existe

ponto que está afastado da primeira bisetriz e, portanto, é possível encontrar valores inesperados ao utilizar o ajuste linear.

Figura 5 – Valores Ajustados x Resíduos

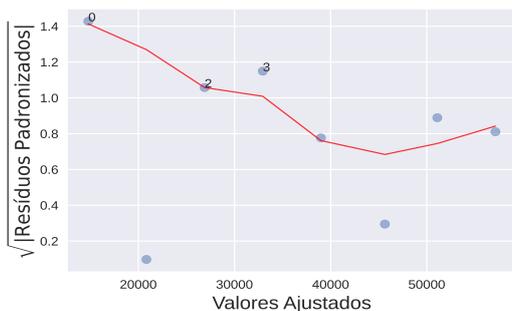


Fonte – Elaborada pelos autores: Código Python 12

Gráfico Escala-Localização

Este gráfico mostra se os resíduos apresentam variância não constante, também conhecida como heterocedasticidade. Quanto mais horizontal for a linha que o representa, mais provável é que os dados sejam homocedásticos. A Figura 6 mostra o gráfico de uma linha não horizontal e, portanto, os dados apresentam variância não constante.

Figura 6 – Valores Escala-Localização



Fonte – Elaborada pelos autores: Código Python 13

Gráfico Alavancagem x Resíduos

Considere uma distribuição $(X, Y) \in \mathbb{R}^2$, com $X = \{x_1, x_2, \dots, x_n\}$ e $Y = \{y_1, y_2, \dots, y_n\}$. Sejam \bar{x} e $\sigma^2(X)$, respectivamente, a média aritmética e a variância de X e definamos

$$H = X(X^t X)^{-1} X^t = \begin{cases} h_{ii} = \frac{1}{n} + \frac{(x_i - \bar{x})^2}{\sigma^2(X)} \\ h_{ij} = \frac{1}{n} + \frac{(x_i - \bar{x})(x_j - \bar{x})}{\sigma^2(X)} \end{cases}$$

a **matriz de projeção**. O elemento h_{ii} é chamado o **ponto de alavancagem** para a i -ésima unidade de dados do conjunto X e mede o quão distante a observação y_i está das demais $n - 1$ observações no espaço definido pelas variáveis independentes, ou seja, mede a influência da i -ésima resposta sobre seu valor ajustado. Como

$$H^2 = [X(X^t X)^{-1} X^t] [X(X^t X)^{-1} X^t] = X I (X^t X)^{-1} X^t = H,$$

ou seja, H é uma matriz idempotente e também é simétrica (definição), temos, ao igualarmos o ii -ésimo elemento de H ao de H^2 , que

$$h_{ii} = h_{ii}^2 + \sum_{i \neq j} h_{ij}^2 \geq 0.$$

Segue que $h_{ii} \geq h_{ii}^2$, implicando em $h_{ii} \leq 1$. Além disso,

$$\begin{aligned} h_{ii} &\geq \sum_{i \neq j} h_{ij}^2 \\ &= \sum_{i \neq j} \left[\frac{1}{n} + \frac{(x_i - \bar{x})(x_j - \bar{x})}{\sigma^2(X)} \right]^2 \\ &\geq \sum_{i \neq j} \left[\frac{1}{n} \right]^2 \\ &\geq \frac{1}{n}. \end{aligned}$$

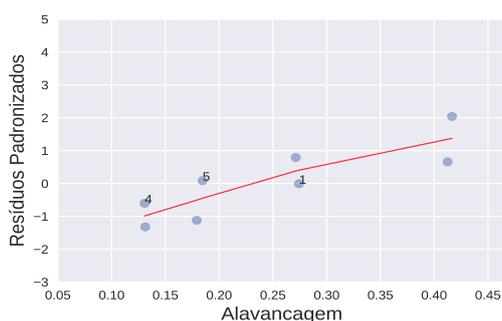
Quanto mais próximo de 1 o valor h_{ii} está, mais atípico é o valor de x_i .

Seja $p = \sum_{i=1}^n h_{ii}$. Se $h_{ii} \geq \frac{2p}{n}$, temos que $y_i \in Y$ é um **ponto de alavancagem**.

Os pontos de alavancagem não são desejados, pois podem causar efeitos prejudiciais no modelo uma vez que os coeficientes da reta de regressão são muito sensíveis a esses pontos. Na prática, pontos que possuem medidas de alavancagem superiores a 0,5 podem ser removidos, desde que sejam poucos se comparados com a quantidade total dos dados.

Observemos que no gráfico de Alavancagem x Resíduos Padronizados (Figura 7), não temos pontos de alavancagem (as alavancagens são inferiores a 0,5).

Figura 7 – Valores Ajustados x Resíduos



Fonte – Elaborada pelos autores: Código Python 14

A análise da regressão apresentada pelos gráficos das figuras 4, 5, 6 e 7, foi conclusiva. Os dados apresentados não possuem uma boa qualidade para que se possa ajustá-los por uma reta. Entretanto, o ajuste linear será utilizado para comparar com resultados obtidos por outros ajustes.

Ajuste Exponencial

Se aplicarmos, no período que compreendeu os anos de 1940 a 2010, um ajuste aos pontos da Tabela 1 utilizando uma curva exponencial

$$P(t) = P_0 e^{\tau t},$$

obteremos $P_0 \simeq 0,00000000004226589$ e $\tau \simeq 0,0173659801642$. Estes valores foram obtidos através do algoritmo 5.

Os valores populacionais obtidos através da equação da curva de ajuste exponencial são apresentados na Tabela 3 e foram determinados utilizando o algoritmo 6.

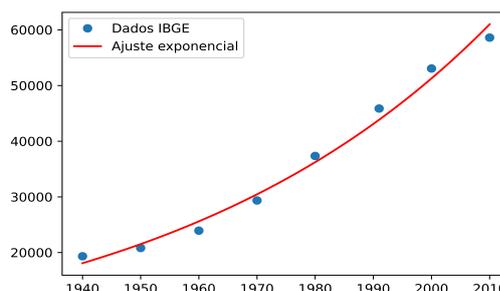
Tabela 3 – População de Cruz das Almas (1940-2010) segundo a curva exponencial

ANO	HABITANTES
1940	18.087
1950	21.517
1960	25.598
1970	30.453
1980	36.228
1991	43.854
2000	51.273
2010	60.996

Fonte – Elaborada pelos autores

Com o intuito de comparar os dados nas tabelas Tabela 1 e Tabela 3, temos que o desvio padrão DP é, aproximadamente, $DP = 1.590$ habitantes. Concluímos, portanto, que a curva exponencial se ajusta melhor à distribuição dada pelo IBGE do que a reta de regressão.

Figura 8 – População de Cruz das Almas (1940-2010) e o ajuste exponencial



Fonte – Elaborada pelos autores: Código Python 8

A Figura 8 ilustra a distribuição dada pelo dados do IBGE (Tabela 1) e a curva de ajuste exponencial.

Com o resultado obtido, podemos estimar que se a população da cidade de Cruz das Almas seguisse um crescimento dado pela curva de ajuste exponencial, em 2020, sua população seria, aproximadamente, de $P(2020) = P_0 \cdot e^{\tau t} = 75.565$ habitantes que, comparado com a população estimada por IBGE (2021), teremos uma diferença de 12.380 habitantes, um erro relativo de, aproximadamente, 19,59%.

Ajuste Sigmoidal

Aplicando um ajuste sigmoidal aos dados apresentados pelo IBGE (Tabela 1) e considerando que esta apresenta um ponto de inflexão tal que $\frac{K}{2} = 45.000$ (dado pela observação do gráfico da Figura 3), os valores dos coeficientes $P_0 \approx 3,35 \cdot 10^{-21}$ e $\eta \approx 0,02942013$ para a Equação 26 (sigmoide) foram determinados utilizando o Código Python em 10.

Os valores populacionais obtidos através da equação da curva de ajuste sigmoidal são apresentados na Tabela 4.

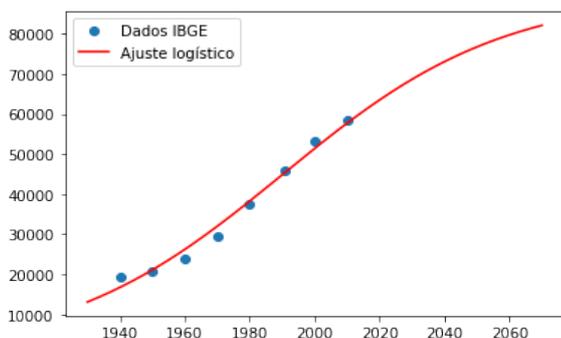
Tabela 4 – População de Cruz das Almas (1940-2010) segundo a curva sigmoidal, com $\kappa = 90.000$

ANO	HABITANTES
1940	16.718
1950	21.096
1960	26.210
1970	31.989
1980	38.278
1991	45.508
2000	51.422
2010	57.729

Fonte – Elaborada pelos autores

Com o intuito de comparar os dados nas Tabela 1 e Tabela 4, temos que o desvio padrão DP é, aproximadamente, $DP = 1.711$ habitantes. Concluímos, portanto, que a curva sigmoidal, com $\kappa = 90.000$, se ajusta melhor à distribuição dada pelo IBGE do que a reta de regressão. Entretanto, seu ajuste não é melhor que o da curva exponencial.

Figura 9 – População de Cruz das Almas (1940-2010) e o ajuste sigmoidal, com $\kappa = 90.000$



Fonte – Elaborada pelos autores: Código Python 10

O gráfico apresentado na Figura 9, ilustra como os dados do IBGE e do ajuste através da curva sigmoide estão próximas. Através da Equação 26), podemos estimar que a população em Cruz das Almas, no ano de 2020, é de 63.535 e se compararmos este valor com o dado divulgado (IBGE, 2021), de 63.591 habitantes, temos um erro absoluto de 56 habitantes (aproximadamente 0,09%).

CONSIDERAÇÕES FINAIS

O objetivo deste trabalho foi o de comparar as estimativas obtidas através dos ajustes linear, exponencial e sigmoidal (logístico) com aquele fornecido pelo IBGE. Para isso foi feito o estudo da dinâmica populacional, apresentado os modelos de Malthus e Verhulst. O ajuste linear foi empregado por ser o modelo mais elementar para o MMQ e os ajustes exponencial e sigmoidal, por serem as soluções dos modelos apresentados por Malthus e por Verhulst, respectivamente. Estipulamos o valor de κ a partir da observação do gráfico da Figura 3, onde, supostamente, há uma inflexão em $\kappa/2 = 45.000$.

Calculando os DP para cada um dos dados dos modelos linear, exponencial e logístico, verificamos que a curva que melhor se ajusta aos dados do IBGE é a exponencial, seguida da sigmoidal (com $\kappa = 90.000$) e do linear.

Ao utilizarmos os modelos linear, exponencial e logístico para estimar a população da cidade de Cruz das Almas em 2020, encontramos 63.185, 75.565 e 63.535 habitantes, respectivamente, enquanto que a predição feita pelo IBGE fora de 63.591 habitantes, apresentando os respectivos erros relativos aproximados de: 0,64%, 19,59% e 0,09%.

Concluímos que, apesar do ajuste ao conjunto de pontos dados pelo IBGE ter sido melhor pela curva exponencial, esta não foi a melhor para estimar o número de habitantes de Cruz das Almas, no ano de 2020. Apesar de não ser a curva sigmoide aquela que melhor se ajusta aos pontos dados na Tabela 1, podemos verificar que esta determinou um valor muito mais próximo ao obtido pelo IBGE.

Referências

BEZERRA, A. D. M. et al. **Equações Diferenciais Aplicadas ao Modelo de Malthus na Dinâmica de Crescimento da População de Bataguassu MS**. *Revista Conexão Eletrônica*, AEMS-MS, v. 13, n. 1, p. 932–946, 2016. Disponível em: <<http://revistaconexao.aems.edu.br/wp-content/plugins/download-attachments/includes/download.php?id=946>>. Citado na página 2.

BOYCE W. E.; DIPRIMA, R. C. **Equações Diferenciais Elementares e Problemas de Valores de Contorno**. Rio de Janeiro: LTC, 2010. Citado 2 vezes nas páginas 3 e 4.

BRASIL. Lei nº 8.184, de 10 de maio de 1991. *Diário Oficial [da] República Federativa do Brasil*, Brasília, DF, 1991. ISSN 1677-7042. Disponível em: <https://www.planalto.gov.br/ccivil_03/leis/l8184.htm>. Citado na página 3.

CUNHA, L. A. G. **A Importância dos Censos**. 2021. Disponível em: <<https://elos.sites.uepg.br/coluna/a-importancia-dos-censos/>>. Acesso em: 08.07.2021. Citado na página 2.

DIAS, V. M.; PESSOA, R. S. **Estudo da Dinâmica de Crescimento Populacional da Cidade de São José dos Campos - SP**. *XX Encontro Latino Americano de Iniciação Científica, XVI Encontro Latino Americano de Pós-Graduação e VI Encontro de Iniciação à Docência*, Universidade do Vale do Paraíba, São José dos Campos - SP, -, n. -, p. -, 2016. Disponível em: <http://cronos.univap.br/cd/INIC_2016/anais/arquivos/RE_1006_0801_01.pdf>. Acesso em: 12.04.2021. Citado na página 2.

HERBÁRIO, U. **Cruz das Almas**. 2021. Disponível em: <<https://ufrb.edu.br/herbario/cruz-das-almas>>. Acesso em: 15.07.2021. Citado na página 2.

HOFFMAN, K.; KUNZE, R. **Álgebra linear**, Tradução de Adalberto Panobianco Bergamasco. 1. ed. São Paulo: Editora Polígono S.A., 1971. Citado na página 7.

IBGE, I. B. d. G. e. E. **Censo demográfico população e habitação**. 1940. 405 p. Disponível em: <https://biblioteca.ibge.gov.br/visualizacao/periodicos/65/cd_1940_p12_t1_ba.pdf>. Acesso em: 24.04.2021. Citado na página 8.

IBGE, I. B. d. G. e. E. **VI Recenseamento Geral do Brasil**. 1952. 15 p. Disponível em: <https://biblioteca.ibge.gov.br/visualizacao/periodicos/70/cd_1950_ba.pdf>. Acesso em: 24.04.2021. Citado na página 8.

IBGE, I. B. d. G. e. E. **Censo demográfico de 1960**. 1960. 80 p. Disponível em: <https://biblioteca.ibge.gov.br/visualizacao/periodicos/68/cd_1960_v1_t8_ba.pdf>. Acesso em: 24.04.2021. Citado na página 8.

IBGE, I. B. d. G. e. E. **VIII Recenseamento geral - 1970**. 1973. 770 p. Disponível em: <https://biblioteca.ibge.gov.br/visualizacao/periodicos/69/cd_1970_v1_t13_ba.pdf>. Acesso em: 24.04.2021. Citado na página 8.

IBGE, I. B. d. G. e. E. **IX Recenseamento geral do Brasil - 1980**. 1983. 222 p. Disponível

em: <https://biblioteca.ibge.gov.br/visualizacao/periodicos/72/cd_1980_v1_t4_n15_ba.pdf>. Acesso em: 24.04.2021. Citado na página 8.

IBGE, I. B. d. G. e. E. **Sinopse preliminar do Censo Demográfico - 1991**. 1991. 48 p. Disponível em: <https://biblioteca.ibge.gov.br/visualizacao/periodicos/309/cd_1991_v6_n15_ba.pdf>. Acesso em: 24.04.2021. Citado na página 8.

IBGE, I. B. d. G. e. E. **Guia do Censo**. 2010. Disponível em: <<https://censo2010.ibge.gov.br/materiais/guia-do-censo/apresentacao.html>>. Acesso em: 09.07.2021. Citado na página 2.

IBGE, I. B. d. G. e. E. **Panorama da cidade de Cruz das Almas**. 2021. Disponível em: <<https://cidades.ibge.gov.br/brasil/ba/cruz-das-almas/panorama>>. Acesso em: 22.07.2021. Citado 4 vezes nas páginas 8, 9, 11 e 12.

MEIRA, J. S. et al. **Modelo malthusiano aplicado ao crescimento populacional do município de Manuel Viana-RS**. *Encontro Regional de Estudantes de Matemática da Região Sul - XX EREMAT*, Universidade Federal do Pampa (UNIPAMPA) - RS, v. 10, n. -, p. -, 2014. Disponível em: <https://eventos.unipampa.edu.br/eremat/files/2014/12/PO_Meira_01550181084.pdf>. Acesso em: 12.04.2021. Citado na página 4.

NUNES, R. R.; TAKAHASHI, L. T. **Dinâmica de Populações: Um Breve Histórico**. 2009. Disponível em: <<http://ime.ufg.br/bienal/2006/poster/roney.pdf>>. Acesso em: 19.11.2020. Citado na página 4.

SDT, S. d. D. T.; MBA, M. d. D. A. **Perfil Territorial - Recôncavo/BA**. 2015. 2 p. Disponível em: <http://sit.mda.gov.br/download/caderno/caderno_territorial_187_Rec%C3%83%C2%B4ncavo%20-%20BA.pdf>. Acesso em: 24.04.2021. Citado na página 8.

SOTOMAYOR T.; MANUEL, J. **Lições de Equações Diferenciais Ordinárias**. Rio de Janeiro: IMPA (Projeto Euclides), 1979. Citado na página 4.

TAVARES, C. A. **O Polo Naval e o município de Rio Grande: Um estudo sobre a dinâmica populacional**. Rio Grande, RS: Universidade Federal do Rio Grande (FURG) - RS, 2017. Monografia. Disponível em: <https://imef.furg.br/images/stories/Monografias/Matematica_aplicada/2017/Cibelle_Tavares.pdf>. Acesso em: 14.08.2021. Citado na página 2.

WETERMAN, D. **Orçamento do Censo diminui 96%**. 2021. Acesso em: 09.07.2021. Citado na página 3.

©2022 by RECET. Este é um artigo de acesso livre, distribuído sob os termos e condições da licença [Creative Commons](https://creativecommons.org/licenses/by-nc-nd/4.0/) (CC BY-NC-ND 4.0).

APÊNDICE A – Códigos Python

Algoritmo 1: Dados tabelados

```
import numpy as np

xk = np.array([1940, 1950, 1960, 1970, 1980, 1991, 2000, 2010])
yk = np.array([19336, 20822, 23941, 29355, 37352, 45874, 53055, 58606])
```

Algoritmo 2: Cálculo dos coeficientes da reta de regressão linear

```
Sxk = xk.sum()
xk2 = xk**2
Sxk2 = xk2.sum()
Syk = yk.sum()
n = len(xk)
prodxkyk = xk*yk
Sprodxkyk = prodxkyk.sum()
alpha1 = (Sxk*Syk-n*Sprodxkyk)/(Sxk**2-n*Sxk2)
alpha2 = (Syk-alpha1*Sxk)/n
print('Os coeficientes são:', '%.7f' % alpha1, '%.3f' % alpha2)
```

Algoritmo 3: Cálculo da população ajustada pela reta de regressão e

```
def r(x):
    return alpha1*x+alpha2
r(xk)
```

Algoritmo 4: Cálculo do desvio padrão

```
DQ = (r(xk)-yk)**2
SDQ = DQ.sum()
DQM = SDQ/n
DP = DQM**(0.5)
```

Algoritmo 5: Cálculo dos coeficientes do ajuste exponencial

```
Sxk = xk.sum()
xk2 = xk**2
Sxk2 = xk2.sum()
Syklog = yklog.sum()
n = len(xk)
prodxkyklog = xk*yklog
Sprdxkyklog = prodxkyklog.sum()
r = (Sxk*Syklog-n*Sprdxkyklog)/(Sxk**2-n*Sxk2)
alpha = (Syklog-r*Sxk)/n
P0 = np.exp(alpha)
print('Coeficientes da curva exponencial:', '%.7f' % r, '%.20f' % P0)
```

Algoritmo 6: Função do ajuste exponencial e os valores ajustados

```
import numpy as np
import matplotlib.pyplot as plt

def func(x):
    return P0*np.exp(r*x)

func(xk)
```

Algoritmo 7: Código Python para o gráfico da reta de regressão

```
import matplotlib.pyplot as plt
plt.scatter(xk,yk,label='Dados IBGE')
p1 = np.polyfit(xk,yk,1)
yr = np.polyval(p1,xk)
plt.plot(xk,yr,color='red',label='Reta de regressão')
plt.legend()
plt.show()
```

Algoritmo 8: Código Python para o gráfico do ajuste exponencial

```
import numpy as np
import matplotlib.pyplot as plt

def func(x):
    return P0*np.exp(beta*x)

xv = np.linspace(1940,2010,1000)

plt.plot(xk, yk, 'o', label='Dados IBGE')
plt.plot(xv, func(xv), 'r-', label='Ajuste exponencial')
plt.legend()
plt.show()
```

Algoritmo 9: Gráfico das soluções do modelo de Verhulst

```
import numpy as np
import matplotlib.pyplot as plt

P1 = 10;P2 = 250;P3 = 700

kappa = 400;eta=0.008

def fver(x,kappa,P0,eta):
    return 1 / (1/kappa + (1/P0 - 1/kappa) *np.exp(-eta * x))

def finvver(x,kappa,P0,eta):
    return -1/eta*np.log((1/x-1/kappa)/(1/P0-1/kappa))

tsob2 = finvver(kappa/2,kappa,P1,eta)

xv = np.linspace(0,800,100)

plt.plot(xv, fver(xv,kappa,P1,eta), 'r-', label='Curva  $\sigma_1$ ')
plt.plot(xv, fver(xv,kappa,P2,eta), 'b-', label='Curva  $\sigma_2$ ')
plt.plot(xv, fver(xv,kappa,P3,eta), 'g-', label='Curva  $\sigma_3$ ')
plt.axhline(y=400,linestyle='--',color='black')
plt.axhline(y=kappa/2,xmin=0,xmax=tsob2/800,linestyle='--',color = (0.211, 0.211, 0.211))
plt.axvline(x=tsob2,ymin=0,ymax=kappa/1400,linestyle='--',color = (0.211, 0.211, 0.211))
plt.xticks([tsob2], ['$t_{\kappa/2}$'])
plt.yticks([P1,P2,P3,kappa,kappa/2], ['$P_1$', '$P_2$', '$P_3$', '$\kappa$', '$\kappa/2$'])
plt.legend()
plt.show()
```

Algoritmo 10: Código Python da sigmoide

```
import numpy as np
import matplotlib.pyplot as plt

tk = xk
n = len(yk)
k = 90000
Yk = []
for i in range(n):
    calc = np.log(yk[i]/(k-yk[i]))
    Yk.append(calc)

Yk = np.array(Yk)

Stk = sum(tk)
tk2 = tk**2
Stk2 = sum(tk2)
SYk = Yk.sum()
prodtkYk = tk*Yk
SprodtkYk = sum(prodtkYk)

alpha1 = (Stk*SYk-n*SprodtkYk)/(Stk**2-n*Stk2)
alpha2 = (SYk-alpha1*Stk)/n
eta = alpha1
P0 = np.exp(alpha2)*k/(1+np.exp(alpha2))

fig = plt.figure()
xv = np.linspace(1930,2070,1000)

def logistic(t):
    return 1/(1/k+(1/P0-1/k)*np.exp(-eta*t))

plt.plot(xk, yk, 'o', label='Dados IBGE')
plt.plot(xv, logistic(xv), 'r-', label='Ajuste logístico')
plt.legend()
plt.show()
```

Algoritmo 11: Resíduos x Valores Ajustados

```
import numpy as np
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
plt.style.use('seaborn') # pretty matplotlib plots
import statsmodels.api as sm
from statsmodels.graphics.gofplots import ProbPlot

plt.rc('font', size=14)
plt.rc('figure', titlesize=18)
plt.rc('axes', labelsiz=15)
plt.rc('axes', titlesize=18)

xk = np.array([1940,1950,1960,1970,1980,1991,2000,2010])
yk = np.array([19336,20822,23941,29355,37352,45874,53055,58606])

X = pd.DataFrame(xk)
y = pd.DataFrame(yk)

model = sm.OLS(y, sm.add_constant(X))
model_fit = model.fit()

model_fitted_y = model_fit.fittedvalues;
model_residuals = model_fit.resid;
model_norm_residuals = model_fit.get_influence().resid_studentized_internal;
model_norm_residuals_abs_sqrt = np.sqrt(np.abs(model_norm_residuals));
model_abs_resid = np.abs(model_residuals);
model_leverage = model_fit.get_influence().hat_matrix_diag;
model_cooks = model_fit.get_influence().cooks_distance[1];

plot_lm_1 = plt.figure();
plot_lm_1.axes[0] = sns.residplot(model_fitted_y,
                                dataframe.columns[1],
                                data=dataframe,
                                lowess=True,
                                scatter_kws={'alpha':0.5},
                                line_kws={'color':'red','lw': 1,'alpha':0.8}
                                );
plot_lm_1.axes[0].set_xlabel('Valores Ajustados');
plot_lm_1.axes[0].set_ylabel('Resíduos');
```

Algoritmo 12: Normal Q-Q

```
QQ = ProbPlot(model_norm_residuals)
plot_lm_2 = QQ.qqplot(line='45', alpha=0.5, color='#4C72B0', lw=1)
plot_lm_2.axes[0].set_xlabel('Quantis teóricos')
plot_lm_2.axes[0].set_ylabel('Resíduos Padronizados');
# annotations
abs_norm_resid = np.flip(np.argsort(np.abs(model_norm_residuals)), 0)
abs_norm_resid_top_3 = abs_norm_resid[:3]
for r, i in enumerate(abs_norm_resid_top_3):
    plot_lm_2.axes[0].annotate(i,
                               xy=(np.flip(QQ.theoretical_quantiles, 0)[r],
                                   model_norm_residuals[i]));
```

Algoritmo 13: Escala-Localização

```
plot_lm_3 = plt.figure();
plt.scatter(model_fitted_y, model_norm_residuals_abs_sqrt, alpha=0.5);
sns.regplot(model_fitted_y, model_norm_residuals_abs_sqrt,
            scatter=False,
            ci=False,
            lowess=True,
            line_kws={'color': 'red', 'lw': 1, 'alpha': 0.8});
#plot_lm_3.axes[0].set_title('Escala-Localização');
plot_lm_3.axes[0].set_xlabel('Valores Ajustados');
plot_lm_3.axes[0].set_ylabel('$\sqrt{|\operatorname{Resíduos} \; \text{Padronizados}|}$');

# annotations
abs_sq_norm_resid = np.flip(np.argsort(model_norm_residuals_abs_sqrt), 0);
abs_sq_norm_resid_top_3 = abs_sq_norm_resid[:3];
for i in abs_norm_resid_top_3:
    plot_lm_3.axes[0].annotate(i,
                              xy=(model_fitted_y[i],
                                  model_norm_residuals_abs_sqrt[i]));
```

Algoritmo 14: Resíduos x Alavancagem

```
plot_lm_4 = plt.figure();
plt.scatter(model_leverage, model_norm_residuals, alpha=0.5);
sns.regplot(model_leverage, model_norm_residuals,
            scatter=False,
            ci=False,
            lowess=True,
            line_kws={'color': 'red', 'lw': 1, 'alpha': 0.8});
plot_lm_4.axes[0].set_xlim(0.05, max(model_leverage)+0.05)
plot_lm_4.axes[0].set_ylim(-3, 5)
plot_lm_4.axes[0].set_xlabel('Alavancagem')
plot_lm_4.axes[0].set_ylabel('Resíduos Padronizados');

leverage_top_3 = np.flip(np.argsort(model_cooks), 0)[:3]
for i in leverage_top_3:
    plot_lm_4.axes[0].annotate(i,
                              xy=(model_leverage[i],
                                  model_norm_residuals[i]));
```
