

Modelagem matemática da incidência de casos de câncer na cidade de Santa Rosa – RS

Mathematical modeling of the incidence of cancer cases in the city of Santa Rosa – RS

DOI:10.34117/bjdv8n7-050

Recebimento dos originais: 23/05/2022

Aceitação para publicação: 30/06/2022

Gilberto Carlos Thomas

Doutor em Engenharia pela Universidade Federal do Rio Grande do Sul (UFRGS)
Instituição: Instituto Federal de Educação, Ciência e Tecnologia Farroupilha (IFFAR)
Endereço: Campus Santa Rosa, Santa Rosa - RS, Brasil
E-mail: gilberto.thomas@iffarroupilha.edu.br

Eduardo Padoin

Doutor em Engenharia Mecânica pela Universidade Federal do Rio Grande do Sul (UFRGS)
Instituição: Instituto Federal de Educação, Ciência e Tecnologia Farroupilha (IFFAR)
Endereço: Campus Santa Rosa, Santa Rosa - RS, Brasil
E-mail: eduardo.padoin@iffarroupilha.edu.br

Pablo Henrique Tolfo

Licenciando em Matemática (IFFAR)
Instituição: Instituto Federal de Educação, Ciência e Tecnologia Farroupilha (IFFAR)
Endereço: Campus Santa Rosa, Santa Rosa - RS, Brasil,
E-mail: pablo.2019002460@aluno.iffar.edu.br

Katieli Graef Ludwig Simionatto

Licencianda em Matemática (IFFAR)
Instituição: Instituto Federal de Educação, Ciência e Tecnologia Farroupilha (IFFAR)
Endereço: Campus Santa Rosa, Santa Rosa - RS, Brasil,
E-mail: katieli.2019002827@aluno.iffar.edu.br

RESUMO

Neste artigo apresenta-se a construção de modelos matemáticos que descrevem o comportamento dos casos de câncer de próstata diagnóstico C61 e mama diagnóstico C50 no município de Santa Rosa – RS no período de 2013 a 2021. Utilizando o Modelo de Malthus, Interpolação polinomial de Lagrange e de Regressão Linear, foram desenvolvidas funções que descrevem o comportamento do número de casos de câncer de próstata e mama. Para obter os dados reais sobre os casos de câncer durante os anos de 2013 até 2021 foi utilizado o DataSUS, o serviço de informações do Sistema Único de Saúde (SUS). Comparando os dados obtidos através dos cálculos e os valores reais foi encontrado um modelo de função, que melhor se adaptou para cada tipo de câncer.

Palavras-chave: câncer, modelagem matemática, Santa Rosa.

ABSTRACT

This article presents the construction of mathematical models that describe the behavior of cases of prostate cancer and breast cancer in the municipality of Santa Rosa - RS from 2013 to 2021. Using the Malthus Model, with polynomial interpolation of Lagrange and of Linear Regression, functions were developed that describe the variation in the number of cases of prostate and breast cancer. DataSUS, the information service of the Unified Health System (SUS) was used to obtain data on cancer cases from 2013 to 2021. Comparing the data obtained through the calculations and the real values, a function model was found, which best adapted to each type of cancer.

Keywords: cancer, mathematical modeling, Santa Rosa.

1 INTRODUÇÃO

A matemática pode ser utilizada como ferramenta para entender e prever fenômenos reais. A modelagem matemática propõe a construção de modelos matemáticos, capazes de descrever, prever e ou resolver problemas reais, baseando-se em dados empíricos e/ou reais. De modo genérico, pode-se dizer que Modelagem Matemática se refere à aplicação da matemática em outras áreas do conhecimento. (BARBOSA, 2004).

O câncer compreende um grupo com mais de 100 doenças que se caracterizam pelo crescimento desordenado de células que podem se espalhar entre os tecidos e órgãos próximos à estrutura afetada inicialmente. Ele é um problema de saúde pública e sua incidência se mostra crescente.

O presente trabalho teve como objetivo analisar o comportamento dos casos de câncer de próstata e mama da cidade de Santa Rosa do Noroeste do Rio Grande do Sul no período de 2013 a 2021. Através do uso da modelagem matemática foi desenvolvido modelos matemáticos que usam dados reais coletados junto do DataSUS, para os casos de câncer de Próstata e de Mama, para os quais foram analisados e a partir destes identificado qual modelo melhor descreveu o comportamento da incidência.

2 METODOLOGIA

Para realizar a pesquisa foram obtidos os dados dos casos de câncer de próstata e mama, considerando os casos por ano da incidência no município de Santa Rosa, RS, através do DataSUS o Departamento de Informática do Sistema Único de Saúde (SUS). Os dados utilizados foram dos anos de 2013 até 2021. Os modelos matemáticos foram desenvolvidos utilizando Modelo de Malthus, interpolação polinomial de Lagrange e de

Ajuste linear Regressão Linear, resultando em modelos para cada tipo de câncer analisado.

Para o câncer de próstata foi considerando o diagnóstico detalhado: C61 - Neoplasia maligna da próstata, que tem como média de aproximadamente 14 casos por ano de residentes do município no período analisado. Já para o câncer de mama foi considerando o diagnóstico detalhado: C50 - Neoplasia maligna da mama, que tem como média de aproximadamente 28 casos por ano de residentes do município no período analisado. A distribuição dos casos de câncer com diagnóstico C61 e C50, residentes no município de Santa Rosa - RS entre o ano de 2013 e 2021, conforme a Tabela I de acordo com dados do DATASUS.

Tabela I: Dados reais do câncer de próstata e mama dos residentes do município de Santa Rosa de 2013 a 2021.

Diagnóstico	Ano									
	2013	2014	2015	2016	2017	2018	2019	2020	2021	
C61	7	17	9	14	11	23	18	19	6	
C50	30	28	30	32	33	28	25	33	17	

Fonte: DATASUS (2022).

Ao descrever a dinâmica populacional em linguagem matemática, Thomas Robert Malthus, em 1798, partiu da ideia de que a variação populacional em relação ao tempo é proporcional ao seu tamanho em cada instante. Desta forma, assumiu-se que existia um coeficiente, de crescimento ou decrescimento calculado a partir da diferença entre a taxa de natalidade e de mortalidade de determinada população.

O método de Lagrange pode ser descrito de forma direta e totalmente analítica, visto que não existem coeficientes desconhecidos, pois os coeficientes da fórmula a partir de Lagrange são os valores de " y_i " dados no problema. A utilização desse método no trabalho utiliza da interpolação polinomial, onde é necessário um ajuste exato aos dados.

A regressão linear encontra a reta que melhor representa as variáveis de entrada com a variável de saída. É uma função linear de primeiro grau para estimar o valor de uma variável dependente y , considerando dados os valores da variável independente x .

3 MODELO DO PROBLEMA

Após a tabulação da coleta de informações do banco de dados do DataSUS, iniciou-se a investigação e aplicação de modelos que poderiam se ajustar com essas informações. Inicialmente, utilizou-se do Modelo de Malthus, na sequência usou-se da interpolação polinomial de Lagrange e da Regressão Linear.

Com intuito de descrever a dinâmica populacional em linguagem matemática, Thomas Robert Malthus, em 1798, partiu do pressuposto de que a variação populacional (dP) em relação ao tempo (dt) é proporcional ao seu tamanho em cada instante (k). Desta forma, assumiu-se que existia um coeficiente k , com crescimento $k>0$ ou decréscimo $k<0$ (EISERMANN; THOMAS, 2020).

$$\frac{dP}{dt} = k P \quad (1)$$

Resolvendo analiticamente (1) e designando c uma constante, obtém-se:

$$\int \frac{dP}{P} = \int k dt \quad (2)$$

$$\ln P = k t + c \quad (3)$$

O processo exponencial em (3), garante que:

$$P = e^{kt+c} \quad (4)$$

$$P = e^{kt} e^c \quad (5)$$

Sendo e^c uma constante de valor desconhecido, ela será denominada de c_1 , levando encontrar a seguinte relação:

$$P(t) = c_1 e^{kt} \quad (6)$$

Denominando P_0 a população no instante inicial ($t = 0$), tem-se que $P_0 = c_1$, e, assim, a solução analítica do Modelo de Malthus torna-se:

$$P_t = P_0 e^{kt} \quad (7)$$

Já, o desenvolvimento do Ajuste de Curva Polinomial, é preciso escolher o menor grau que atenda a todas restrições, com aproximação aceitável dos dados calculados. Se

por acaso o polinômio gerado no ajuste de curvas for de grau $(n - 1)$, ele será idêntico ao polinômio gerado na interpolação. Diante disso, os parâmetros são obtidos de acordo com Barroso, *et al.* (1987) solucionando o sistema em forma matricial de ordem n , que pode ser escrito como:

$$\left\{ \begin{array}{l} (n + 1)a_0 + a_1 \sum_{i=0}^n x_i + a_2 \sum_{i=0}^n x_i^2 + \dots + \sum_{i=0}^n x_i^n = \sum_{i=0}^n y_i \\ a_0 \sum_{i=0}^n x_i + a_1 \sum_{i=0}^n x_i^2 + a_2 \sum_{i=0}^n x_i^3 + \dots + \sum_{i=0}^n x_i^{n+1} = \sum_{i=0}^n x_i \cdot y_i \\ a_0 \sum_{i=0}^n x_i^2 + a_1 \sum_{i=0}^n x_i^3 + a_2 \sum_{i=0}^n x_i^4 + \dots + \sum_{i=0}^n x_i^{n+2} = \sum_{i=0}^n x_i^2 \cdot y_i \\ \vdots \\ a_0 \sum_{i=0}^n x_i^n + a_1 \sum_{i=0}^n x_i^{n+2} + a_2 \sum_{i=0}^n x_i^{n+2} + \dots + \sum_{i=0}^n x_i^{2n} = \sum_{i=0}^n x_i^n \cdot y_i \end{array} \right. \quad (8)$$

Entretanto, a maneira utilizada nesse trabalho para resolver o problema da interpolação polinomial é realizada através dos polinômios de Lagrange. A partir de um conjunto de pontos $\{x_j\}_{j=1}^n$ dados dois a dois, é definido os polinômios de Lagrange como os polinômios de grau $n - 1$ que satisfazem:

$$L_k(x_j) = \begin{cases} 1, & \text{se } k = j \\ 0, & \text{se } k \neq j \end{cases} \quad (9)$$

Assim, o polinômio $p(x)$ de grau $n - 1$ que interpola os pontos dados, isto é, $p(x_j) = y_j, j = 1, \dots, n$ é dado por:

$$p(x) = y_1 L_1(x) + y_2 L_2(x) + \dots + y_n L_n(x) = \sum_{k=1}^n y_k L_k(x) \quad (10)$$

Para construir os polinômios de Lagrange, podemos analisar sua forma fatorada, ou seja:

$$L_k(x) = c_k \prod_{\substack{j=1 \\ j \neq i}}^n (x - x_j) \quad (11)$$

Onde o coeficiente c_k é obtido da condição $L_k(x_k) = 1$:

$$L_k(x_k) = c_k \prod_{\substack{j=1 \\ j \neq i}}^n (x_k - x_j) \Rightarrow c_k = \frac{1}{\prod_{\substack{j=1 \\ j \neq i}}^n (x_k - x_j)} \quad (12)$$

Portanto, substituindo c_k (12) em (11) e $L_k(x)$ (11) em (10), temos que a fórmula (13) da interpolação lagrangeana.

$$L_i(x) = \sum_{i=0}^n y_i \cdot \prod_{\substack{j=1 \\ j \neq i}}^n \frac{(x - x_j)}{(x_i - x_j)} \quad (13)$$

O modelo de Regressão Linear consiste de uma variável chamada de dependente estar relacionada a uma ou mais variáveis independentes por uma equação linear. Assim, a linha de regressão minimiza os desvios quadrados dos dados reais. Para se obter o cálculo da equação da reta basta aplicar a seguinte equação:

$$y = a + bx \quad (14)$$

Para se encontrar os valores de a e b , segue a fórmula conforme descreve Gaither e Frazier (2006):

$$a = \frac{\sum x^2 \sum y - \sum x \sum xy}{n \sum x^2 - (\sum x)^2} \quad (15)$$

$$b = \frac{n \sum xy - \sum x \sum y}{n \sum x^2 - (\sum x)^2} \quad (16)$$

De acordo com as equações (15) e (16) apresentadas, o cálculo dos coeficientes a e b tem finalidade minimizar a soma dos desvios quadrados dos dados reais da linha do gráfico. A variável n corresponde o período considerado no estudo.

4 RESULTADOS E DISCUSSÕES

De acordo com os três métodos investigados nessa pesquisa em relação ao câncer de próstata, foram substituídos os dados conforme informações da DataSUS para a cidade de Santa Rosa - RS no período analisado. Para tanto, a modelagem realizada do Modelo de Malthus, da interpolação polinomial de Lagrange de ordem oito e da Regressão Linear, são apresentadas respectivamente nas funções abaixo (17), (18) e (19), onde $P(t)$ representa a incidência de casos de câncer de Próstata C61, e t é o tempo em anos. Os resultados obtidos em cada modelo servem para representar o comportamento dos dados reais do diagnóstico de câncer C61 do ano de 2013 a 2021.

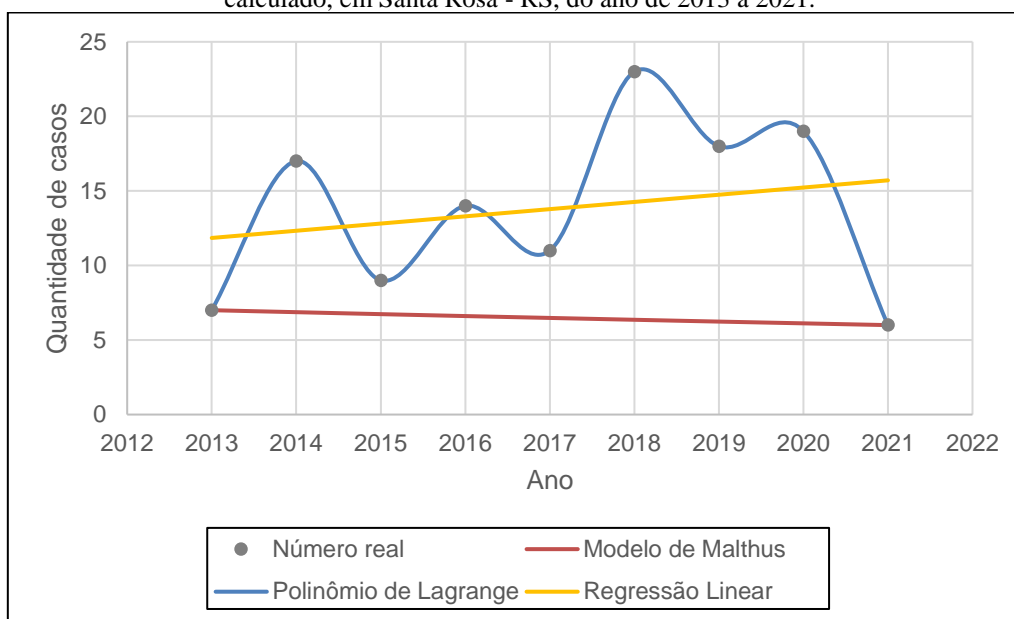
$$P(t) = 7e^{-0,019268834t} \quad (17)$$

$$P(t) = -\frac{821}{40320}t^8 + \frac{1171}{1440}t^7 - \frac{39263}{2880}t^6 + \frac{22439}{180}t^5 - \frac{3892507}{5760}t^4 + \frac{3177409}{1440}t^3 - \frac{42271127}{10080}t^2 + \frac{501359}{120}t - 1620 \quad (18)$$

$$P(t) = 0,4833t + 11,361 \quad (19)$$

Com esses dados calculados, utilizou-se os modelos para plotar os resultados conforme a gráfico 1, para melhor análise do comportamento das funções, ao longo do período de tempo da pesquisa. Os dados reais do DataSUS são número real representados por pontos e as funções são representados com linhas sólidas.

Gráfico 1: Comportamento do número de casos do diagnóstico C61 - Neoplasia maligna da próstata real e calculado, em Santa Rosa - RS, do ano de 2013 a 2021.



Fonte: Dados da Pesquisa (2022).

Desse modo, analisando graficamente o comportamento geral dos pontos percebe-se que o Modelo de Malthus e de Regressão Linear não tiveram uma projeção muito próxima dos dados reais, com exceção do primeiro e último ponto no de Malthus, em que devem ser exatamente os mesmos. Já a interpolação polinomial de Lagrange apresentou um desempenho muito bom correlacionando os dados reais e os modelados. Logo, concluiu-se que para a modelagem do câncer de próstata, o Modelo de Lagrange descreveu melhor o comportamento dos dados reais, uma vez que não se trata de um comportamento com crescimento exponencial ou linear.

Na sequência, utilizando esse mesmo período e métodos, foi aplicado à relação ao câncer de mama, sendo substituídos os dados conforme informações da DataSUS para a cidade de Santa Rosa/RS. A modelagem realizada do Modelo de Malthus, da interpolação polinomial de Lagrange de ordem oito e da Regressão Linear, são apresentadas respectivamente nas funções (20), (21) e (22), onde $M(t)$ representa a incidência de casos de câncer de mama C50, e t é o tempo em anos.. Os resultados obtidos em cada modelo servem para representar o comportamento dos dados reais do diagnóstico de câncer C50 do ano de 2013 a 2021.

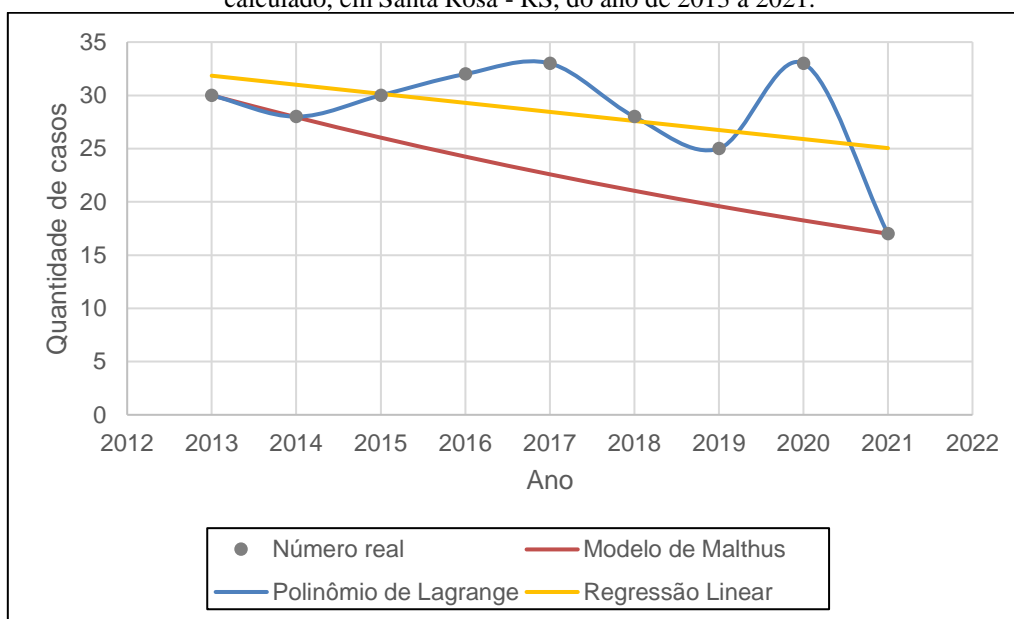
$$M(t) = 30e^{-0,070998004t} \tag{20}$$

$$M(t) = \frac{7}{5760}t^8 - \frac{547}{10080}t^7 + \frac{571}{576}t^6 - \frac{6953}{720}t^5 + \frac{315223}{5760}t^4 - \frac{265429}{1440}t^3 + \frac{11497}{32}t^2 + \frac{308249}{840}t + 176 \tag{21}$$

$$M(t) = -0,85t + 32,694 \tag{22}$$

No Gráfico 2, apresenta-se os resultados dos modelos para o câncer de Mama C50 em comparação aos dados reais(número real) de câncer de mama.

Gráfico 2: Comportamento do número de casos do diagnóstico C50 - Neoplasia maligna da mama real e calculado, em Santa Rosa - RS, do ano de 2013 a 2021.



Fonte: Dados da Pesquisa (2022).

Percebe-se pelo gráfico 2 que a interpolação polinomial de Lagrange teve uma aproximação muito boa comparada com os dados reais, dentre os outros dois métodos utilizados. Dessa forma, no caso da interpolação de Lagrange, o polinômio de ordem oito foi eficiente ao descrever o comportamento dos números de casos com diagnóstico de ambos os tipos de câncer abrangidos nessa pesquisa. Logo, em um primeiro momento de análise gráfica o polinômio de Lagrange é o mais indicado, pois se obteve maior efetividade na representação das informações reais comparado com os outros modelos.

Segundo Damião *et al.* (2015, p.81), “Estimou-se 68.800 casos novos de câncer de próstata em 2014. Esses valores correspondem a um risco de aproximadamente 70,42 casos novos a cada 100 mil homens.” Santa Rosa possui de acordo com o Data Sebrae aproximadamente 77.200 habitantes destes 38.357 são homens, caso a média brasileira se aplicasse seriam 27,01 novos casos todos os anos. Podemos ver que os números de novos casos de câncer de próstata são abaixo da média brasileira.

Já os números de câncer de mama têm estimativa menor, como mostram os dados do Instituto Nacional de Câncer:

O Instituto Nacional de Câncer (INCA) estima que para cada ano do triênio 2020/2022, sejam diagnosticados no Brasil 66.280 novos casos de câncer de mama, com um risco estimado de 61,61 casos a cada 100 mil mulheres (INSTITUTO NACIONAL DO CÂNCER, 2020).

Santa Rosa possui aproximadamente 38.843 mulheres, caso a média brasileira se aplicasse seriam 23,95 novos casos de câncer de mama todos os anos. Porém a média de novos casos é de 28,44 por ano, sendo assim um pouco acima da média nacional.

O câncer de mama e de próstata possuem um fator hereditário. Essa pode ser uma explicação para as diferenças da média brasileira, quanto mais casos existem mais podem vir a ocorrer.

Como explica BARROS *et al.* (2001) sobre o câncer de mama, de 90% a 95% dos casos de câncer de mama são esporádicos e 5% a 10% são hereditários. E também DAMIÃO *et al.* (2015) sobre o câncer de próstata, caso um parente de 1º grau tenha a doença a chance é 2 vezes maior de acometer a pessoa.

Também existem outros fatores que aumentam as chances de ser acometido pelos cânceres de próstata e mama, como a idade e hábitos alimentares e físicos. Na idade quanto mais velho a pessoa, maior a chance de contrair um dos cânceres.

Os modelos usando o Método de Malthus, Interpolação de Lagrange e Regressão Linear permitiu ajustar 6 funções que descrevem o comportamento da incidência de câncer de próstata do diagnóstico C61 e mama diagnóstico C50 no período de 2013 a 2021, do município de Santa Rosa - RS. Sendo três funções para cada tipo de câncer.

Com a finalidade de verificar o grau de relação entre as variáveis reais e as encontradas pelo Modelo de Malthus, Interpolação Polinomial de Lagrange e Regressão Linear, utilizou-se do cálculo do coeficiente de Pearson (STANTON, 2001). Este coeficiente, de acordo com Filho e Junior (2009), trata-se de uma medida de associação

linear entre variáveis, a qual busca mensurar a direção e o grau da relação linear entre duas variáveis quantitativas, conforme fórmula:

$$r = \frac{1}{n-1} \sum \left(\frac{x_i - \bar{X}}{S_x} \right) \left(\frac{y_i - \bar{Y}}{S_y} \right) \quad (23)$$

Os cálculos dos coeficientes de correlação de Pearson foram realizados através da ferramenta do *Excel*, de modo que o processo seja facilitado e agilizado, visto a prontidão da ferramenta tecnológica. Na tabela abaixo, apresentamos os resultados de tal coeficiente para os modelos utilizados no presente trabalho, os quais podem exprimir valores entre -1 a 1. Quando o coeficiente de correlação está próximo de zero indica que não existe relação entre as variáveis analisadas, mas quando se aproxima dos extremos, -1 e 1, mais forte é a relação.

Tabela II: Coeficientes de Correlação de Pearson em relação aos 3 modelos.

Correlação de Pearson	Modelo de Malthus	Polinômio de Lagrange	Regressão Linear
Câncer de Próstata	-0,235709453	1	0,224722941
Câncer de Mama	0,422309605	1	0,462999062

Dados da Pesquisa (2022).

Conforme a tabulação dos cálculos do Coeficiente de Person para os métodos utilizados na pesquisa, a interpolação polinomial de Lagrange de ordem oito demonstrou ser mais próximo dos dados reais. Apresentando uma correlação muito boa diante do câncer de próstata e câncer de mama, que implica em uma relação mais fidedigna aos valores reais e o referido modelo, no caso da correlação igual a 1, significa que as variáveis estão diretamente correlacionadas. Os Modelos de Malthus e de Regressão Linear demonstraram estar mais longe da realidade e, portanto, o polinômio de Lagrange é o mais indicado para o uso nessa situação, até mesmo, para a aplicação de projeções futuras, em razão de seu menor índice de erro relativo conforme o Coeficiente de Correlação de Pearson.

5 CONCLUSÕES

Pode-se destacar nesta pesquisa a importância da modelagem matemática em situações reais, em que um modelo pode se comportar melhor que outro dependendo da situação e dos dados reais disponibilizados. Assim, com os modelos construídos podem

ajudar a realizar futuras projeções em vista do número de casos novos para câncer de próstata e mama para residentes do município Santa Rosa/RS.

Os cânceres de próstata e mama possuem campanhas de prevenção que ocorrem todos os anos, porém mesmo ocorrendo todos os anos, observa-se uma variação de incidência no número de diagnósticos. Este é um dos motivos pelos quais a formulação de modelos exatos torna-se difícil de ser realizado.

Sendo assim, com os dados coletados durante os anos de 2013 a 2021 a respeito do número de diagnóstico de câncer dos tipos C61 - Neoplasia maligna da próstata e C50 - Neoplasia maligna da mama na cidade de Santa Rosa/RS, foram possíveis a elaboração de modelos matemáticos que descrevessem a realidade do momento da pesquisa. Bem como, através desses modelos foi possível analisar qual se enquadra melhor para demonstrar o comportamento real destas doenças crônicas na cidade de Santa Rosa - RS, em que o polinômio de Lagrange se mostrou mais eficaz para descrever a realidade de ambos os diagnósticos de câncer analisados, com erro dos dados modelados comprovado a partir correlação de Pearson.

Portanto, nota-se nesse trabalho que o uso da modelagem matemática que geram funções que representam situações reais, do cotidiano da sociedade, é viável e possível de serem feitas, como a feita usando a Interpolação polinomial de Lagrange. O uso dos dados reais atrelados a modelagem matemática, permitem que haja o fortalecimento dos conceitos matemáticos aproximados com a realidade. Assim, entende-se a modelagem como instrumento fundamental para que seja realizado as análises temporais da sociedade, especialmente voltado para o uso como recurso despertador de ações políticas e beneficentes à população.

REFERÊNCIAS

BARROS A. C.S. D. *et al.* **Diagnóstico e tratamento do câncer de mama.** In: Projeto diretrizes. Brasília: Associação Médica Brasileira/Conselho Federal de Medicina; 2001. v. 2.

BARBOSA, J. C. **Modelagem Matemática: O que é? Por que? Como?.** Veritati, n. 4, p. 73- 80, 2004.

BARROSO, L. C. *et al.* **Cálculo Numérico (com aplicações)- 2º edição.** Editora Harbra Ltda, São Paulo, 1987.

Brasil, Ministério da Saúde. **Banco de dados do Sistema Único de Saúde-DATASUS, Painel-Oncologia.** Disponível em <http://tabnet.datasus.gov.br/cgi/dhdat.exe?PAINEL_ONCO/PAINEL_ONCOLOGIABR.def>. Acesso em: 24/04/2022.

DAMIÃO, Ronaldo et al. **Câncer de próstata. Revista Hospital Universitário Pedro Ernesto (TÍTULO NÃO-CORRENTE),** [S.l.], v. 14, ago. 2015. ISSN 1983-2567.

EISERMANN, Jonatan Ismael, THOMAS, Gilberto Carlos, **Equações Diferenciais Aplicadas à Dinâmica Populacional do Município de Santa Rosa/RS,** Brazilian Journal of Development, Curitiba, v. 6, n. 7, p. 44612-44616, jul. 2020. ISSN 2525-8761

GAITHER, N.; FRAZIER, G. **Administração da produção e operações.** 8. ed. São Paulo: Thomson Learning, 2006.

INSTITUTO NACIONAL DO CÂNCER (Brasil). Estômago. In: INSTITUTO NACIONAL DO CÂNCER (Brasil). **Síntese de Resultados e Comentários.** [Brasília, DF]: Instituto Nacional do Câncer, 2020. Disponível em: [https://www.inca.gov.br/estimativa/sintese-de-resultados-e-comentarios#:~:text=Para%20o%20Brasil%2C%20estimam%2Dse,mil%20mulheres%20\(Tabela%201\)](https://www.inca.gov.br/estimativa/sintese-de-resultados-e-comentarios#:~:text=Para%20o%20Brasil%2C%20estimam%2Dse,mil%20mulheres%20(Tabela%201).). Acesso em: 25/04/2022.

RODRIGUES, Sandra Cristina Antunes. **Modelo de Regressão linear e suas Aplicações.** 2012. Disponível em: <<http://hdl.handle.net/10400.6/1869>>. Acesso em: 20/04/2022.

STANTON, J. M. **Galton, Pearson, and the peas: A brief history of linear regression for statistics instructors.** Journal of Statistical Education, vol.9, n.3, 2001.