



COINTER PDVAgro 2020

V CONGRESSO INTERNACIONAL DAS CIÊNCIAS AGRÁRIAS

Edição 100% virtual | 02 a 05 de dezembro

ISSN:2526-7701 | PREFIXO DOI:10.31692/2526-7701

MODELOS DE REGRESSÃO DO NÚMERO DE CASOS DE LEPTOSPIROSE NO RECIFE/PE, BRASIL

MODELOS DE REGRESIÓN DEL NÚMERO DE CASOS DE LEPTOSPIROSIS EN RECIFE/PE, BRASIL

REGRESSION MODELS THE OF NUMBER CASES OF LEPTOSPIROSIS IN RECIFE/PE, BRAZIL

Apresentação: Comunicação Oral

Ruben Vivaldi Silva Pessoa¹; Jucarlos Rufino de Freitas²; Leika Irabele Tenório de Santana³; Mickaelle Maria de Almeida Pereira⁴; Moacyr Cunha Filho⁵

DOI: <https://doi.org/10.31692/2526-7701.VCOINTERPDVAgro.0247>

RESUMO

A Leptospirose é uma doença de caráter zoonótico que afeta os animais domésticos, silvestres e o homem relacionando-se, principalmente, a variáveis ambientais e socioeconômica. Desta forma, o presente estudo teve por objetivo investigar qual o modelo de regressão com melhor ajuste para os dados dos casos de leptospirose no município de Recife/PE, fazendo uso das técnicas ligadas para ajuste e seleção de modelos. Os dados presentes nesse estudo são decorrentes do total de casos mensais de leptospirose em Recife/PE, sendo obtidos por meio do Sistema Nacional de Agravos de Notificação (SINAN) e dos dados de precipitação pluviométrica, através da Agência Pernambucana de Águas e Clima (APAC), para o período de janeiro de 2007 a dezembro de 2019. Os modelos regressão linear simples é um método que estabelece uma relação linear, descrevendo as interações entre duas variáveis dependente (resposta) e independente (entrada). Para obtenção de um bom modelo necessita-se estimar seus parâmetros, assim é realizado o método de mínimos quadrados, que buscar minimizar a soma dos quadrados dos erros. Para a seleção do melhor modelo foi utilizado o critério *Akaike Information Criteria* (AIC) e pôr fim a análise de resíduos e diagnostico foi utilizada para investigar a adequabilidade de um modelo de regressão com base nos resíduos. A análise apresentou uma forte correlação entre a precipitação acumulada mensal dos postos 30 e 265 com o número de casos de leptospirose em Recife/PE, sendo de aproximadamente 0.747 e de aproximadamente 0.679, para o posto 30 e 265, respectivamente. Através da análise de resíduos, verifica-se que os dados se ajustam bem ao modelo encontrado. As técnicas usadas propiciaram uma análise associativa entre as variáveis em estudo,

¹ Pós Graduação em Biometria e Estatística Aplicada, Universidade Federal Rural de Pernambuco, rvsp1@de.ufpe.br

² Pós Graduação em Biometria e Estatística Aplicada, Universidade Federal Rural de Pernambuco, jucarlos123@hotmail.com

³ Pós Graduação em Biometria e Estatística Aplicada, Universidade Federal Rural de Pernambuco, leikatenorio10@gmail.com

⁴ Pós Graduação em Biometria e Estatística Aplicada, Universidade Federal Rural de Pernambuco, mickaellealmeida1@gmail.com

⁵ Doutor, Universidade Federal Rural de Pernambuco, moacyr.cunhafo@ufrpe.br

MODELOS DE REGRESSÃO DO NÚMERO DE CASOS DE LEPTOSPIROSE NO

propiciando uma compreensão acerca de como se comportam esses dados ao longo dos anos nessa região, além de uma eventual estimativa de como tais dados se apresentarão nos próximos anos em Recife.

Palavras-Chave: Precipitação, zoonose, análise de regressão.

RESUMEN

La leptospirosis es una enfermedad zoonótica que afecta a los animales domésticos y salvajes y a los seres humanos y se relaciona principalmente con variables ambientales y socioeconómicas. Así pues, el presente estudio tenía por objeto investigar qué modelo de regresión se ajustaba mejor a los datos de los casos de leptospirosis en el municipio de Recife/PE, utilizando las técnicas vinculadas para el ajuste y la selección del modelo. Los datos presentes en este estudio se derivan del total de casos mensuales de leptospirosis en Recife/PE, y se obtienen a través del Sistema Nacional de Notificación Agravos (SINAN) y los datos de precipitaciones, a través de la Agencia del Agua y el Clima de Pernambuco (APAC), para el período comprendido entre enero de 2007 y diciembre de 2019. Los modelos de regresión lineal simple son un método que establece una relación lineal, describiendo las interacciones entre dos variables dependientes (respuesta) e independientes (entrada). Para obtener un buen modelo es necesario estimar sus parámetros, por lo que se realiza el método de los mínimos cuadrados, que busca minimizar la suma de los cuadrados de los errores. Para seleccionar el mejor modelo se utilizó el Criterio de Información de Akaike (AIC) y para poner fin al análisis de residuos y el diagnóstico se utilizó para investigar la idoneidad de un modelo de regresión basado en los residuos. El análisis mostró una fuerte correlación entre la precipitación mensual acumulada de los puestos 30 y 265 con el número de casos de leptospirosis en Recife/PE, siendo aproximadamente 0,747 y aproximadamente 0,679 para los puestos 30 y 265, respectivamente. A través del análisis de los residuos, se verifica que los datos encajan bien con el modelo encontrado. Las técnicas utilizadas proporcionaron un análisis asociativo entre las variables en estudio, lo que permitió comprender cómo se comportan esos datos a lo largo de los años en esta región, así como una eventual estimación de cómo se presentarán esos datos en los próximos años en Recife.

Palabras Clave: Precipitación, zoonosis, análisis de regresión.

ABSTRACT

Leptospirosis is a zoonotic disease that affects domestic animals, wild animals and humans relating mainly to environmental and socioeconomic variables. Thus, the present study aimed to investigate which regression model best fits the data of leptospirosis cases in the municipality of Recife/PE, making use of the linked techniques for model adjustment and selection. The data present in this study are from the total of monthly cases of leptospirosis in Recife/PE, being obtained through the National System of Agravos Notification (SINAN) and the pluviometric precipitation data, through the Pernambuco Water and Climate Agency (APAC), for the period of January 2007 to December 2019. Simple linear regression models are a method that establishes a linear relationship, describing the interactions between two variables dependent (response) and independent (input). To obtain a good model it is necessary to estimate its parameters, so the method of least squares is performed, which seeks to minimize the sum of the squares of errors. To select the best model the Akaike Information Criteria (AIC) was used and to put an end to the waste analysis and diagnosis was used to investigate the suitability of a regression model based on the waste. The analysis showed a strong correlation between the monthly accumulated precipitation of posts 30 and 265 with the number of leptospirosis cases in Recife/PE, being approximately 0.747 and approximately 0.679 for post 30 and 265, respectively. Through the analysis of residues, it is verified that the data fit well with the model found. The techniques used provided an associative analysis among the variables under study, providing an understanding of how these data behave over the years in this region, as well as an eventual estimate of how such data will present themselves in the coming years in Recife.

Keywords: Precipitation, zoonosis, regression analysis.

INTRODUÇÃO

A leptospirose se apresenta como uma doença de origem muito antiga, em que uma das primeiras aparições de organismos como forma de espiroqueta, bactéria presente na manifestação da doença, foi notada no início do século XX, por meio de uma investigação de tecidos de pacientes diagnosticados com febre amarela (STIMSON, 1907). Apesar disso, os estudos iniciais deste gênero estiveram limitados a poucos institutos de pesquisa, devido ao fato de bactérias do gênero *Leptospira* possuírem um crescimento dependente de nutrientes especiais em seu meio de cultura, o que exige meios de cultura alternativos para nutrição (PAPLOSKI et.al., 2013).

Presente em diversos países, a leptospirose atinge tanto animais como humanos, e nos casos mais sérios, pode gerar infecção grave. Os animais silvestres e domésticos funcionam como hospedeiros sendo os roedores, morcegos e marsupiais considerados como os principais hospedeiros. O ser humano é tido como hospedeiro acidental e terminal dentro da cadeia de transmissão da patologia (BRASIL, 2014).

No Brasil, a leptospirose é uma doença que ocorre habitualmente e com incidência significativa em dada população e/ou região, apresentando uma média de aproximadamente 13.000 notificações de casos por ano, em que 3.500 destes são confirmados e a taxa de letalidade média é de 10,8%. Em território nacional, alguns surtos epidêmicos surgem em decorrência de questões ambientais como períodos chuvosos associados a ocorrência de enchentes e inundações. A maior parte da população brasileira afetada pela doença é do sexo masculino, que habitam nas zonas urbanas e periurbanas, apresentando baixos níveis de escolaridade e faixa etária economicamente ativa (PEREIRA, 2013).

Em Pernambuco, de 2008 a 2017, foram constatados 1170 casos de leptospirose, o que correspondeu a 3,46% do total do país para o período em questão, onde Recife, como esperado, dentre todos os municípios de Pernambuco, apresentou o maior número de casos (560), aproximadamente 47,86% dos casos de todo o estado, tendo o município vizinho de Olinda com o segundo maior número de casos, mas com menos de 100 notificações, apontando que Recife contém um total de ocorrências de Leptospirose bem maior em relação aos demais municípios do estado (DANTAS, 2018).

No período estudado, o ano com índices mais altos de notificações de casos de leptospirose em Recife foi o de 2011, e ao analisar o comportamento do número de casos nesse ano, pode-se concluir que a distribuição dos casos ao longo dos meses não é uniforme, sinalizando para a presença de surtos epidêmicos, principalmente nos meses de fortes chuvas,

MODELOS DE REGRESSÃO DO NÚMERO DE CASOS DE LEPTOSPIROSE NO

onde há enchentes e inundações.

Considerando o conjunto de variáveis que podem influenciar na incidência de casos de leptospirose numa determinada região, esse estudo teve por objetivo investigar qual o modelo de regressão com melhor ajuste para os dados dos casos de leptospirose no município de Recife/PE, fazendo uso das técnicas ligadas para ajuste e seleção de modelos.

FUNDAMENTAÇÃO TEÓRICA

Aqui, encontram-se descritas, as suposições para o modelo de regressão linear, as técnicas estatísticas utilizadas para ajuste de modelos e para estimativa dos parâmetros a eles associados, além dos critérios de seleção e análise de resíduos e diagnóstico do modelo.

Modelo

Nos modelos de regressão linear simples, o valor da variável sob estudo ou variável resposta Y é expressa por uma equação linear de apenas uma variável explicativa denominada X_1 , em que para X_1 tem-se um coeficiente linear β_1 associado, como a seguir (MONTGOMERY; PECK; VINING, 2001):

$$Y_i = \beta_0 + \beta_1 X + \epsilon_i \text{ para } i = 1, 2, \dots, n \quad (1)$$

em que β_0 representa o intercepto do modelo, β_1 o parâmetro da inclinação da regressão linear denominado coeficiente angular, e ϵ representa o erro aleatório do modelo.

O modelo de Regressão Linear Simples supõe que o resíduo $\epsilon_i = Y_i - \hat{Y}_i$ segue distribuição Normal com média zero e variância constante σ^2 , ou seja, $\epsilon_i \sim N(0, \sigma^2)$. E, para um conjunto de observações considera-se erros independentes e identicamente distribuídos.

A estimação dos parâmetros desse modelo é realizada pelo método de mínimos quadrados, no qual busca-se minimizar

$$\sum_{i=1}^n \epsilon_i^2 = \sum_{i=1}^n [Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_i]^2 \quad (2)$$

em relação aos parâmetros β_0 e β_1 . Dessa forma, obtém se

[4]

$$\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{x} \quad (3)$$

em que $X = \frac{\sum_{i=1}^n x_i}{n}$ e $Y = \frac{\sum_{i=1}^n y_i}{n}$ são as médias da variável explicativa e da variável dependente, respectivamente.

Também, após minimizar a soma dos quadrados dos desvios ϵ , se tem:

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n x_i Y_i - n \bar{x} \bar{Y}}{\sum_{i=1}^n x_i^2 - n \bar{x}^2}. \quad (4)$$

Desse modo, o modelo de regressão linear simples é definido por

$$\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 x \quad (5)$$

em que \hat{Y} é o estimador pontual da média da variável Y para um determinado valor de x, ou seja,

$$\hat{Y} = E(\widehat{Y|x_i}) = \hat{\beta}_0 + \hat{\beta}_1 x, \text{ para } i = 1, \dots, n.$$

Vale salientar que ainda se faz necessária a verificação da adequabilidade do modelo e a autenticidade das suposições feitas para o modelo de regressão linear simples, dentre as quais se destacam:

- (i) O modelo é linear nos parâmetros;
- (ii) A esperança dos erros é 0;
- (iii) Os erros são homoscedásticos ou têm $\text{Var}(\epsilon_i) = \sigma^2$ constante;
- (iv) Os erros são não correlacionados, ou seja, ϵ_i e ϵ_j são independentes ($i \neq j$);
- (v) Os erros seguem distribuição Normal, $\epsilon_i \sim N(0, \sigma^2)$.

A regressão linear pode ser generalizada de tal forma que se tenha uma regressão polinomial. Para isso, pode-se considerar a regressão linear simples como regressão polinomial de um polinômio de grau um. Assim, para regressão polinomial, ao invés de se ajustar a eq. (1), utiliza-se:

$$Y_i = \beta_0 + \beta_1 X_1 + \beta_2 X_2^2 + \dots + \beta_m X_m^2 + \epsilon_i \quad (6)$$

MODELOS DE REGRESSÃO DO NÚMERO DE CASOS DE LEPTOSPIROSE NO

Para os modelos polinomiais, independente da ordem do polinômio, as variáveis explicativas devem ser quantitativas. O principal problema com o uso de modelos de regressão é com a extrapolação, caso em que os valores da variável resposta se encontram fora de um intervalo, a partir do conhecimento do comportamento de tal variável dentro desse intervalo.

Critério de seleção do modelo

Para a seleção do melhor modelo foi utilizado o critério *Akaike Information Criteria* (AIC) proposto por Akaike (1974), dado por:

$$AIC = \ln(\hat{\sigma}_\varepsilon^2) \frac{2(p + q)}{n} \quad (7)$$

em que levam em conta a variância do erro, o tamanho da amostra e os valores dos parâmetros estimados, sendo importantes para verificar o ajuste do modelo com base no fenômeno analisado.

Análise de resíduos e diagnóstico do modelo

Análise dos resíduos e diagnóstico do modelo são um conjunto de técnicas estatísticas utilizadas para investigar a adequabilidade de um modelo de regressão com base nos resíduos, sendo observada as seguintes suposições:

- (i) Independência;
- (ii) Normalidade;
- (iii) Homocedasticidade;
- (iv) Não existência de *outliers* (PAULA, 2004).

METODOLOGIA

O presente estudo é de natureza quantitativa, sendo identificado como estudo epidemiológico observacional descritivo, de onde é feita a análise dos dados da leptospirose na população pertencente ao município de Recife/PE. Então, essa pesquisa busca observar o comportamento da doença no município, investigando a presença e dimensão da relação entre

o total de casos notificados e variáveis importantes como o período.

Os dados presentes nesse estudo são decorrentes do total de casos mensais de leptospirose em Recife/PE, sendo obtidos por meio do Sistema Nacional de Agravos de Notificação (SINAN) e dos dados de precipitação pluviométrica, através da Agência Pernambucana de Águas e Clima (APAC), para o período de janeiro de 2007 a dezembro de 2019.

Inicialmente, os dados apresentavam 10 postos de precipitações, mas para oito destes postos, o acesso aos dados não foi disponível, tornando a análise desses dados inexecutável. Dentre os postos onde se foi possível inferir algo (Posto 30 e Posto 265), buscou-se analisar a relação entre o total de casos de leptospirose por mês e os índices pluviométricos registrados nos postos presentes.

Para a análise estatística dos dados, foi utilizado o software estatístico *R Core Team*.

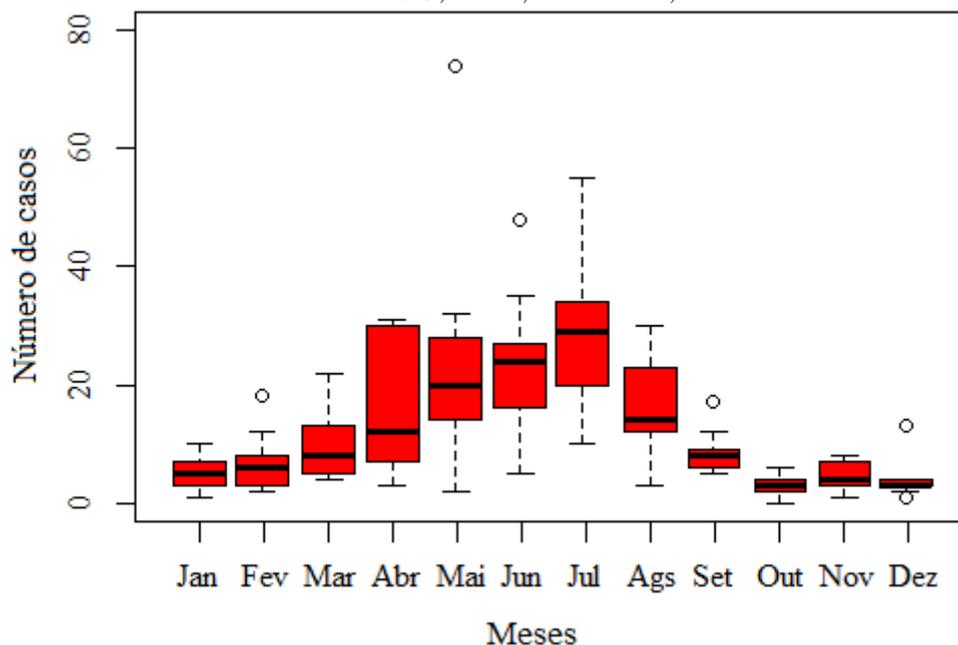
RESULTADOS E DISCUSSÃO

Na Figura 1, pode-se observar que o número de casos de leptospirose tende a ser maior no período entre os meses de Abril e Agosto, apresentando mais notificações em Julho. Isso ocorre em função dos meses serem marcados pelas fortes chuvas na região, o que provoca um aumento no total de notificações de leptospirose no município de Recife. Nota-se a existência de alguns *outliers* no boxplot do número de casos mensais de leptospirose, em particular nos meses de maio e junho, coincidindo com os meses mais chuvosos no município. A amplitude para os meses de maio a agosto são maiores em relação os outros meses, mostrando assim alta variação nos casos mensais de leptospirose. A partir das medianas fornecido pelo boxplot, observa-se que os dados são assimétricos, tanto positivamente, como negativamente.

Segundo Barcellos et al. (2003), no estudo distribuição espacial da leptospirose no Rio Grande do Sul, demonstram a mesma característica observada que os casos de leptospirose indicam uma variabilidade e forte assimetria com média 12,5 casos por 100 mil habitantes e desvio padrão de 49,9. Resultado similar em Dozsa, Del Monego e Kummer (2016), demonstrando uma leve assimetria no mesmo fenômeno.

MODELOS DE REGRESSÃO DO NÚMERO DE CASOS DE LEPTOSPIROSE NO

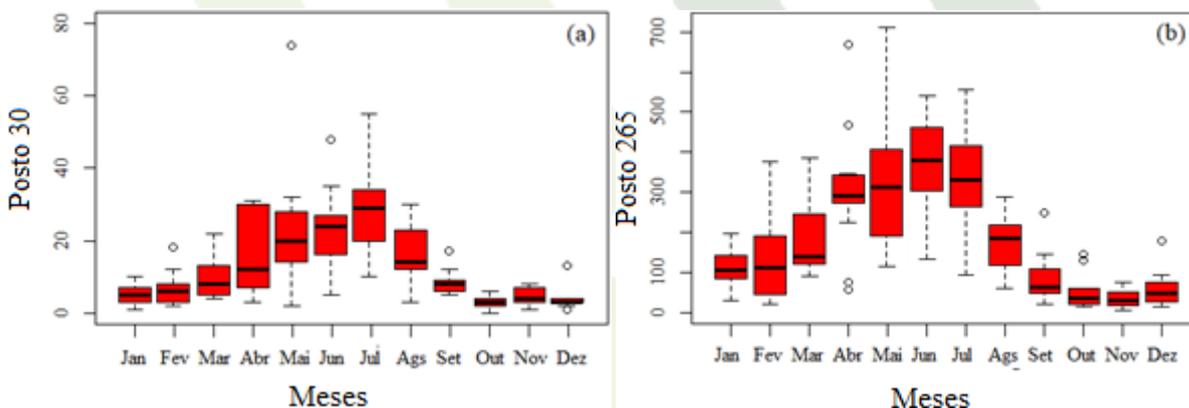
Figura 1. Boxplot do número de casos mensais acumulados de Leptospirose, no período de janeiro de 2007 a dezembro de 2019, Recife, Pernambuco, Brasil.



Fonte: Própria (2020).

Ainda acerca da Figura 2, é possível notar que os meses de maio, junho e julho apresentam os maiores índices pluviométricos no posto 30 e 265 no município de Recife/PE. Esse fato ocorre em função desses meses estarem próximos do inverno no hemisfério Sul, fase em que as chuvas no Recife costumam ser mais frequentes. Vale dizer ainda que existe uma forte tendência que nos meses em que as chuvas se tornam mais frequentes, aliado a presença das inundações, os casos de leptospirose no município do Recife possam aumentar.

Figura 2. Boxplot da precipitação acumulada posto 30 (a) e posto (265), no período de janeiro de 2007 a dezembro de 2019, Recife, Pernambuco, Brasil.



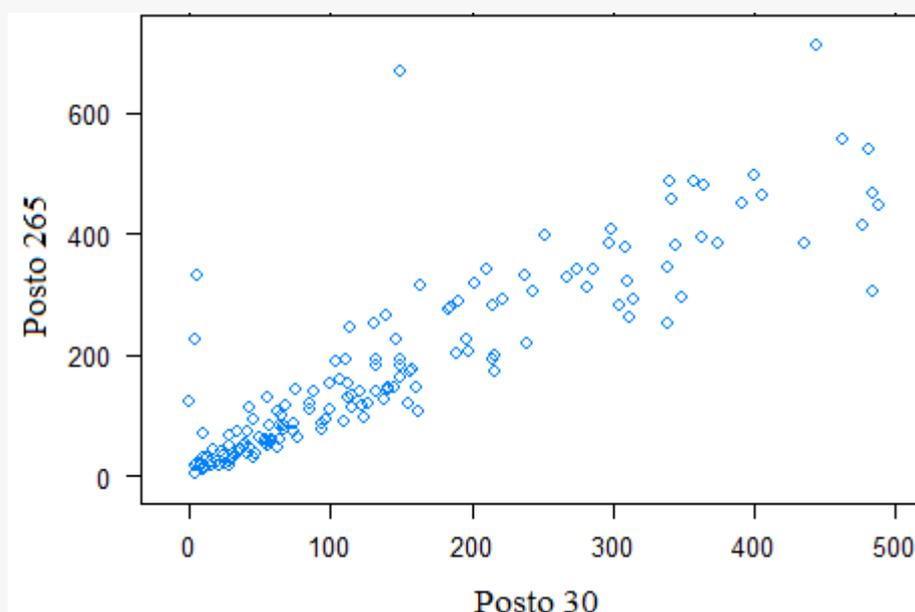
Fonte: Própria (2020).

A análise apresentou uma forte correlação entre a precipitação acumulada mensal dos postos 30 e 265 com o número de casos de leptospirose em Recife/PE, sendo de

aproximadamente 0.747 para o posto 30 e de aproximadamente 0.679 para o posto 265. Diante disso, se faz necessário investigar, de diferentes formas, a relação existente entre os casos de leptospirose e os índices pluviométricos para o município no período em questão.

No ajuste inicial do modelo de regressão, para um modelo com a variável resposta sendo o total de casos de leptospirose e os dois postos considerados como variáveis explicativas, percebe-se que essas variáveis apresentaram alto nível de correlação, aproximadamente 0.88 (Figura 3), o que sinaliza a presença de multicolinearidade no ajuste desse modelo (DAOUD, 2017), o que inviabiliza a análise de regressão, sendo necessário recorrer a um ajuste de um novo modelo com diferentes variáveis que explique melhor os dados aqui presentes.

Figura 3. Correlação entre número de casos mensais e precipitação posto 30 (a) e posto 265 (b), no período de janeiro de 2007 a dezembro de 2019, Recife, Pernambuco, Brasil.



Fonte: Própria (2020).

Diante disso, foi proposto um modelo de regressão linear simples em que a variável resposta segue se referindo ao número de casos mensais de leptospirose e a variável explicativa agora é a soma dos índices nos dois postos (Tabela 1). Para esse modelo, o intercepto não se mostrou significativo ao nível de significância 5%, mas a 10%.

Tabela 1. Estimativa dos parâmetros significativos do modelo de regressão linear dos casos de leptospirose inserindo o somatório dos postos (30 e 265), no período de janeiro de 2007 a dezembro de 2019, Recife, Pernambuco, Brasil.

Coefficientes	Estimativa	Std. Erro	p-valor
Intercept	1,971	1,009	0,052

MODELOS DE REGRESSÃO DO NÚMERO DE CASOS DE LEPTOSPIROSE NO

Posto soma	0,032	0,002	< 0,01
R^2	0,544	R^2_{ajust}	0,541

Fonte: Própria (2020).

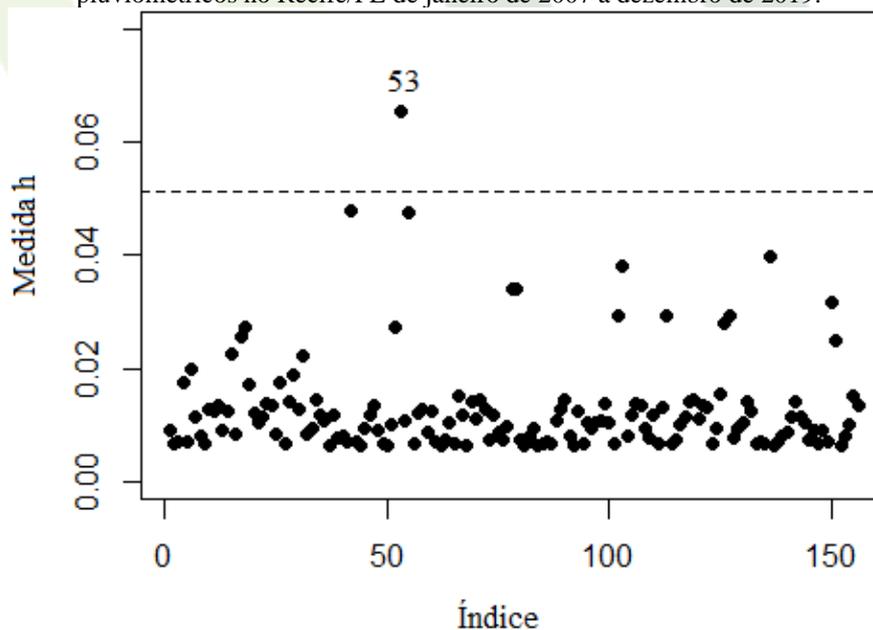
Assim, o modelo ajustado foi representado pela seguinte equação: $Tot_{letp} = 1,971 + 0.032 * Posto\ soma$. Após a construção do modelo se faz necessário realizar a análise de resíduos e diagnóstico do modelo.

Observa-se que os dados se ajustam bem ao modelo encontrado, mas que algumas observações se apresentam como atípicas, sendo em sua maioria, justificadas pelo período de chuvas, em que a disseminação da doença aumenta, bem como o acontecimento de surtos num município (Silva; Conceição; Neto, 2019).

Após o ajuste do modelo, se faz necessário observar se existem falhas no ajuste do modelo proposto. Como, por exemplo, multicolinearidade e heterocedasticidade. Que podem ser facilmente detectados a partir da análise de resíduos e diagnóstico (Figura 4-7).

Na Figura 4, os valores da medida h aparecem plotados versus os índices. Analisando este gráfico, pode-se perceber que o ponto 53 se distanciam dos demais pontos, sendo considerada um ponto de alavancagem no modelo. Contudo, um estudo extensivo realizado por Paula (2004), demonstram que a retirada de uma observação ou um conjunto de observações podem trazer erros no ajuste.

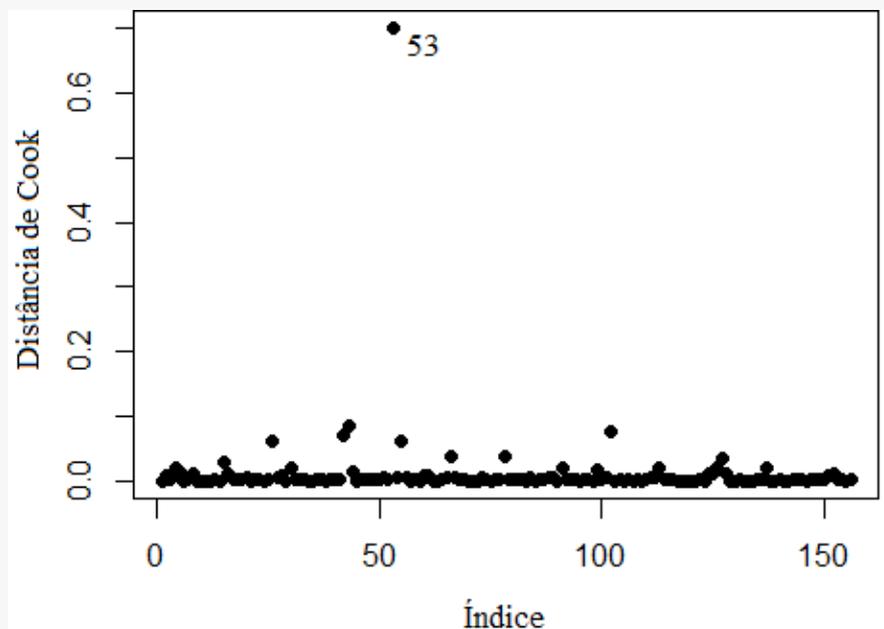
Figura 4: Medida h de diagnóstico do modelo referente aos dados dos casos de leptospirose e dos índices pluviométricos no Recife/PE de janeiro de 2007 a dezembro de 2019.



Fonte: Própria (2020).

Na Figura 5, apresenta o gráfico de pontos influentes, observa-se que todos pontos se encontram abaixo de 1, indicando não ter pontos aberrantes no modelo. Contudo, uma das observações chama a atenção: a 53, próximo de 0.7. Essa observação é de maio de 2011 e a soma dos índices pluviométricos foi de 1155,8 mm, e apresentando 74 casos de leptospirose. Tais valores, por estarem bem acima da média fazem com que essa observação possa ser considerada como influente dentro do conjunto de observações que possui um modelo ajustado, respectivamente.

Figura 5: Distância de cook do diagnóstico do modelo referente aos dados dos casos de leptospirose e dos índices pluviométricos no Recife/PE de janeiro de 2007 a dezembro de 2019.

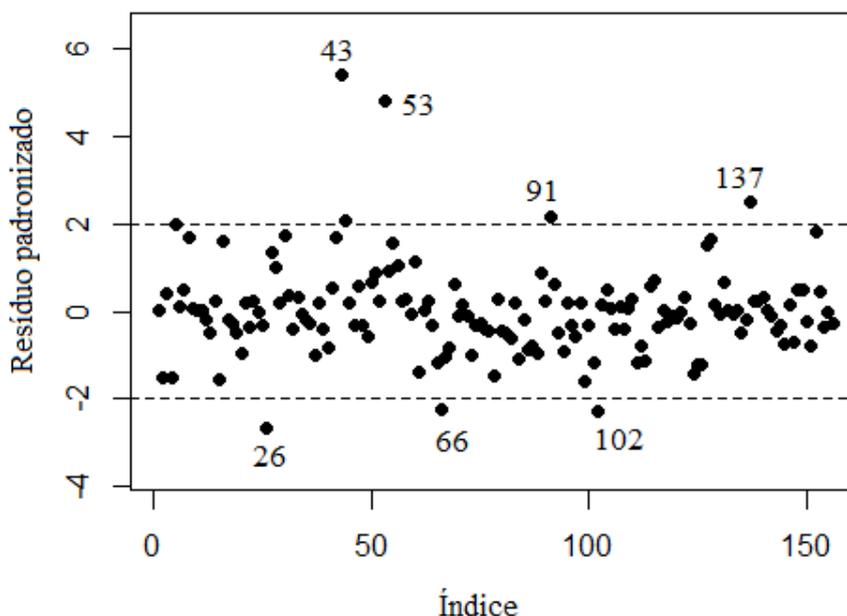


Fonte: Própria (2020).

Analisando os resíduos padronizados (Figura 6), observa-se que os resíduos apresentam um comportamento aleatório. Isso aponta homocedasticidade no modelo, o que significa que a variância dos resíduos é constante ao longo das observações. No entanto, constata-se que existem sete observações que se distanciam dos limites estipulados entre eles a observação 53 com observada nsa figuras acima.

MODELOS DE REGRESSÃO DO NÚMERO DE CASOS DE LEPTOSPIROSE NO

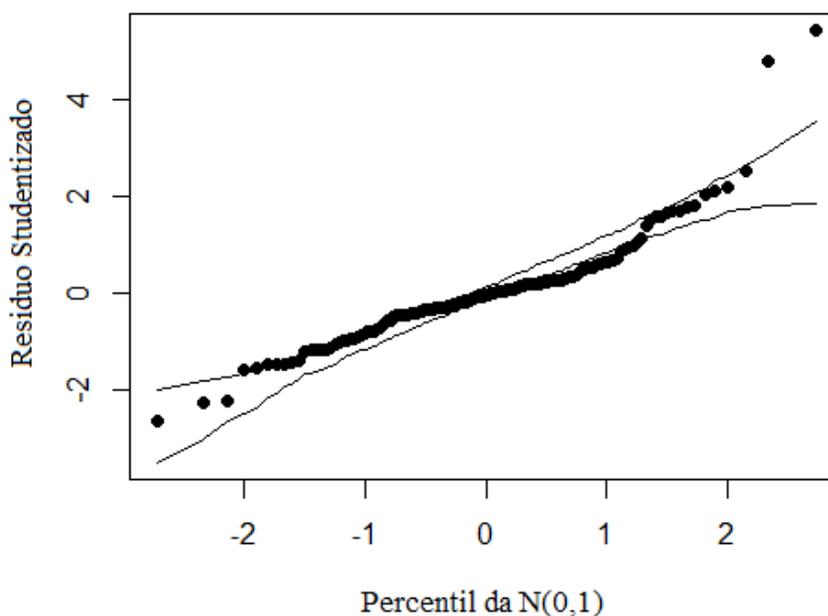
Figura 6: Resíduo padronizado versus índice do modelo referente aos dados dos casos de leptospirose e dos índices pluviométricos no Recife/PE de janeiro de 2007 a dezembro de 2019.



Fonte: Própria (2020).

Na Figura 7, apresentad o gráfico do percentil versus o resíduo studentizado, nota-se que a maioria das observações esta bem comportada com apenas duas observações distantes do envelope do modelo ajustado, sendo duas possíveis observações que interferem nas demais. Assim, de acordo com o grafico, há fortes sinais de que os resíduos seguem uma distribuição normal de probabilidade.

Figura 7: Percentil da distribuição Normal versus o resíduo padronizado do modelo referente aos dados dos casos de leptospirose e dos índices pluviométricos no Recife/PE de janeiro de 2007 a dezembro de 2019.



Fonte: Própria (2020).

CONCLUSÕES

As ferramentas utilizadas nesse estudo viabilizaram uma análise associativa entre as variáveis presentes no banco de dados, possibilitando um entendimento acerca de como se comportam esses dados ao longo dos anos nessa região, além de uma eventual estimativa de como tais dados se apresentarão nos próximos anos em Recife, sendo necessário respeitar as condições para as aplicações de cada técnica. Para a sociedade, tudo isso é de grande representatividade, a fim de que a população possa estar sempre prevenida e atenta aos cuidados necessários, especialmente nos períodos dos surtos epidêmicos, e também para que os governantes junto as secretarias e agencias possam atentar para a necessidade de projetos que visem a prevenção e o tratamento da população.

AGRADECIMENTO

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) – Código de Financiamento 001. À Universidade Federal Rural de Pernambuco (UFRPE); ao Programa de Pós-Graduação em Biometria e Estatística Aplicada (PPGBEA) e Sistema Nacional de Agravos de Notificação (SINAN).

REFERÊNCIAS

AKAIKE, H. **A new look at the statistical model identification.** IEEE transactions on automatic control, v. 19, n. 6, p. 716-723, 1974.

BARCELLOS, C. et al. **Distribuição espacial da leptospirose no Rio Grande do Sul, Brasil: recuperando a ecologia dos estudos ecológicos.** Cadernos de Saúde Pública, v. 19, p. 1283-1292, 2003.

Brasil. Ministério da Saúde (MS). Secretaria de Vigilância em Saúde. Departamento de Vigilância das Doenças Transmissíveis. **Leptospirose: diagnóstico e manejo clínico.** Brasília: MS; 2014.

DANTAS, Elias Ferraz. Redes neurais artificiais aplicadas à previsão de surtos de leptospirose. 2018. 50 f. **Trabalho de Conclusão de Curso** (Bacharelado em Sistemas de Informação) – Unidade Acadêmica de Serra Talhada, Universidade Federal Rural de Pernambuco, Serra Talhada, 2018.

MODELOS DE REGRESSÃO DO NÚMERO DE CASOS DE LEPTOSPIROSE NO

DAOUD, J. I. **Multicollinearity and regression analysis**. Journal of Physics: Conference Series. IOP Publishing, p. 012009, 2007.

DOZSA, B.; DEL MONEGO, M. L. C.; KUMMER, L. **Modelagem geostatística da ocorrência de casos de leptospirose e alagamentos no município de Curitiba-PR, no ano de 2014**. Holos, v. 1, p. 381-393, 2016.

MONTGOMERY, D.C.; PECK, E. A.; VINING, G. G. **Introduction to linear regression analysis**. John Wiley & Sons, p.13, 3rd edition, 2001.

PAPLOSKI, I. A. D. et al. História natural da leptospirose urbana: influência do sexo e da idade no risco de infecção, progressão clínica da doença e óbito. 2013. **Tese de Doutorado**. Centro de Pesquisas Gonçalo Moniz.

PAULA, G. A. **Modelos de regressão: com apoio computacional**. São Paulo: IME-USP, p. 28-55, 2004.

PEREIRA, C.A.R. Custo social da leptospirose no Brasil e o efeito de chuvas extremas em Nova Friburgo para o incremento de casos da doença. 2013 **Dissertação**. Rio de Janeiro. Escola Nacional de Saúde Pública Sergio Arouca, 2013.

R CORE TEAM. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>, 2020.

SILVA, A. E. P.; DE SOUZA C., G. M.; NETO, F. C. **Distribuição espacial da leptospirose em santa catarina nos anos de 2008 e 2011**. Geosaude-2019. 2019.

STIMSON, A. M. **Note on an organism found in Yellow-Fever tissue**. Public Health Reports, v.22, n.18. 1907.