

**Universidade de São Paulo  
Instituto de Matemática e Estatística**

**Centro de Estatística Aplicada**

**Relatório de Análise Estatística**

RAE-CEA-24P29

**RELATÓRIO DE ANÁLISE ESTATÍSTICA SOBRE O PROJETO:**

**“Análise do perfil de atendimentos nas unidades de pronto atendimento e  
notificações de dengue”**

**Fábio Mateus Souza  
Victor Lima de Oliveira  
Airlane Pereira Alencar**

**São Paulo, novembro de 2024**

**CENTRO DE ESTATÍSTICA APLICADA - CEA – USP**

**TÍTULO:** Relatório de Análise Estatística sobre o Projeto: “Análise do perfil de atendimentos nas unidades de pronto atendimento e notificações de dengue”.

**PESQUISADOR(A):** Vanessa Coelho de Aquino Benjaino Ferraz

**ORIENTADOR(A):** Profa. Dra. Airlane Pereira Alencar

**INSTITUIÇÃO:** Centro de Informações Estratégias de Vigilância em Saúde – Campo Grande (CIEVS-CG)

**FINALIDADE DO PROJETO:** Publicação, Outra – Melhoria de processo de trabalho

**RESPONSÁVEIS PELA ANÁLISE:** Fábio Mateus Souza

Victor Lima de Oliveira

Airlane Pereira Alencar

**REFERÊNCIA DESTE TRABALHO:** SOUZA, F.M.; OLIVEIRA, V.L.; ALENCAR, A.P.

**Relatório de análise estatística sobre o projeto: “Análise do perfil de atendimentos nas unidades de pronto atendimento e notificações de dengue”.**

São Paulo, IME-USP, 2024. (RAE–CEA-24P29)

## FICHA TÉCNICA

### REFERÊNCIAS BIBLIOGRÁFICAS:

- Bussab, W.O. e Morettin, P.A. (2012). **Estatística Básica**. São Paulo: Ed. Saraiva, 7ª ed. 539 p.
- Box, G. E., e Jenkins, G.M., (1994). **Time Series Analysis: Forecasting and Control**. 5ª ed., Wiley, Nova Jersey.
- Gabriel, A. F. B., Alencar, A. P. e Miraglia, S. G. E. K., (2019). **Dengue outbreaks: unpredictable incidence time series**. *Epidemiology and Infection* 147, e116, 1-7, disponível em <https://doi.org/10.1017/S0950268819000311>
- Paula, G. A. (2004). **Modelos de regressão: com apoio computacional** . São Paulo: IME-USP. Acesso em: 26 nov. 2024, disponível em [https://www.ime.usp.br/~giapaula/texto\\_2024.pdf](https://www.ime.usp.br/~giapaula/texto_2024.pdf)

### PROGRAMAS COMPUTACIONAIS UTILIZADOS:

Microsoft Excel for Windows (versão 2016)

Rstudio (versão 2023.10.31)

### TÉCNICAS ESTATÍSTICAS UTILIZADAS

Análise de Dados Categorizados (06:030)

Séries Temporais (11:010)

### ÁREA DE APLICAÇÃO

Médica (14:040)

## Resumo

Neste estudo foi realizada a análise de dados de prontuários eletrônicos de atendimentos realizados em unidades de saúde em Campo Grande – MS, cujos pacientes foram diagnosticados com suspeita de dengue. Também foi analisado um banco de dados de notificações de dengue em Campo Grande, proveniente do Sistema de Informação de Agravos de Notificação (SINAN), órgão responsável pelos dados epidemiológicos de doenças de notificação compulsória no Brasil.

A análise descritiva realizada no estudo identificou padrões importantes sobre os dados de notificações. Variáveis como sexo, idade, sintomas, comorbidades e evolução foram exploradas. A análise dos sintomas revelou alta prevalência de febre, mialgia e cefaleia (acima de 77%). Comorbidades foram pouco relatadas, com hipertensão (2,4%) e diabetes (1%) como as mais frequentes, o que sugere possível subnotificação. Quanto à evolução dos casos, 99,2% resultaram em cura, e apenas 0,2% apresentaram óbito.

Na regressão logística, foi avaliado o impacto de variáveis preditoras na classificação dos casos como dengue ou descartados. Resultados indicaram que sintomas como febre, mialgia, cefaleia e exantema aumentam significativamente as chances de diagnóstico positivo para dengue. Comorbidades e outros fatores também foram considerados no modelo, mostrando um ajuste de alta qualidade e corroborando a importância dessas variáveis para a identificação e manejo clínico dos casos de dengue.

Além disso, avaliou-se a relação temporal entre contagens de atendimentos e de notificações, para verificar se os atendimentos poderiam prever a curva de notificações. Foi utilizada a técnica de correlação cruzada entre as séries temporais, e a forte autocorrelação no lag 0 indicou que não há defasagem temporal entre as séries, permitindo focar na previsão direta da curva de notificações. Ajustando o modelo SARIMA para séries temporais, foi possível obter um modelo razoável para previsões da série de notificações, mesmo considerando o histórico de alta volatilidade da mesma.

## **Sumário**

|                                         |           |
|-----------------------------------------|-----------|
| <b>1. Introdução</b>                    | <b>7</b>  |
| <b>2. Objetivo</b>                      | <b>7</b>  |
| <b>3. Descrição do estudo</b>           | <b>8</b>  |
| <b>4. Descrição das variáveis</b>       | <b>8</b>  |
| <b>5. Análise descritiva</b>            | <b>9</b>  |
| <b>6. Modelo de regressão logística</b> | <b>13</b> |
| <b>7. Análise de séries temporais</b>   | <b>14</b> |
| <b>8. Conclusões</b>                    | <b>17</b> |
| <b>APÊNDICE A</b>                       | <b>19</b> |
| <b>APÊNDICE B</b>                       | <b>26</b> |

## **1. Introdução**

O Centro de Informações Estratégicas de Vigilância em Saúde de Campo Grande (CIEVS-CG) é uma unidade regional da Rede Nacional de Monitoramento e Respostas às Emergências em Saúde Pública da Secretaria de Vigilância em Saúde (SVS/MS). Suas atividades são voltadas para detecção e rápida disseminação de informações que proporcionem resposta aos riscos de disseminação de doenças/agravs, visando apoiar a intervenção precoce e oportuna sobre emergências em saúde pública.

Dentro das análises epidemiológicas realizadas pela instituição, está o monitoramento dos casos de dengue na cidade de Campo Grande, realizado através de análises de bancos de dados de atendimento e de notificação dos casos da doença. Durante a consulta do paciente em uma unidade de pronto atendimento, os profissionais são orientados a realizar uma notificação ao SINAN sempre que houver um caso de suspeita de dengue. O SINAN armazena os dados de notificações enviados de unidades ao longo de todo o território brasileiro, e disponibiliza aos órgãos competentes para auxílio nas tomadas de decisão.

## **2. Objetivos**

O projeto tem como objetivo analisar a relação entre os dados de atendimento dos pacientes diagnosticados ou com suspeita de dengue, e os dados de notificações de dengue enviados ao Sistema de Informação de Agravos de Notificação (SINAN). Outra questão a ser verificada é o nível de gravidade da dengue associada ao perfil do paciente e suas características, como presença de comorbidades, sexo, idade, entre outras.

### 3. Descrição do estudo

O estudo foi realizado com base em um banco de dados referente aos atendimentos dos pacientes para os quais houve suspeita de dengue em algumas unidades de pronto atendimento de Campo Grande - MS, selecionadas a critério dos pesquisadores. Foram excluídos do estudo os atendimentos e notificações os quais se tratavam de revisita do paciente, menos de 7 dias após a primeira consulta, na qual ocorreu a suspeita da doença. Esse foi implementado verificando, para cada atendimento, se havia outro registro do mesmo paciente nos 7 dias anteriores. Essa condição foi avaliada por meio de uma operação lógica envolvendo a data do atendimento e os identificadores dos pacientes, eliminando registros que não atendiam ao critério. A abordagem garantiu a exclusão de visitas repetidas dentro do intervalo estipulado, totalizando 41.370 atendimentos e 278.832 notificações restantes.

### 4. Descrição das variáveis

As variáveis presentes no estudo foram:

- Sexo: masculino e feminino
- Idade: idade do paciente na data da consulta em anos
- Data notificação: data que o atendimento de dengue foi notificado
- Sintomas: sintomas apresentados pelo paciente durante o atendimento
- Comorbidades: comorbidades que o paciente possui no momento do atendimento

- Sorotipo: identificação do tipo de dengue do paciente, DENV-1, DENV-2, DENV-3 e DENV-4
- Hospitalização: indicador se o paciente foi internado ou não
- Classificação: dengue, descartado, não identificado, dengue alarme, dengue grave
- Evolução: cura, ignorado, óbito por outras causas, óbito investigação, óbito pelo agravado (dengue).

## 5. Análise descritiva

Nesta seção, será abordada a qualidade dos dados, incluindo a verificação de dados faltantes, a análise dos principais indicadores descritivos e a distribuição das variáveis de interesse.

### Verificação de Dados Faltantes

Conforme ilustrado na Tabela A.1, as variáveis de data de notificação, sexo e idade apresentam preenchimento completo, com 100% de dados válidos. No entanto, observamos que a variável raça e classificação possuem uma proporção de dados faltantes inferior a 2,2%. A variável evolução tem 17,6% de dados faltantes, o que sugere um padrão de subnotificação ou falha de preenchimento em alguns atendimentos.

As variáveis de sintomas e comorbidades compartilham uma taxa significativa de dados faltantes, correspondente a 47,3%. Essa alta proporção sugere que o preenchimento dessas informações pode ter começado em uma data posterior ao início da coleta de dados, já que o banco abrange registros desde 2008. A variável sorotipo, com 99% de dados faltantes, será desconsiderada nas análises subsequentes, dado o baixo número de informações disponíveis.

### **Distribuição Temporal dos Dados Faltantes**

A Tabela A.2, que mostra as variáveis por ano e percentual de dados faltantes por ano, revela um padrão interessante: a partir de 2017, as variáveis de sintomas, comorbidades e evolução apresentam uma taxa de preenchimento satisfatória, com menos de 5% de dados faltantes. Dada essa melhoria na qualidade dos dados, optou-se por restringir as análises aos registros a partir de 2017, o que resulta em um total de 55 mil registros para a análise descritiva.

### **Idade**

A análise da variável idade apresentou uma média de 31 anos e uma mediana de 29. Conforme histograma na Figura B.1 (Bussab e Morettin, 2012), a distribuição das idades se aproxima de uma distribuição normal, com o primeiro quartil (Q1) em 17 anos e o terceiro quartil (Q3) em 45 anos, o que reflete a diversidade etária da população atendida.

### **Sexo**

A distribuição da variável sexo está dentro do esperado para a população em análise. Conforme a Figura B.2, 55% dos atendimentos foram de pacientes do sexo feminino, enquanto 45% foram de pacientes do sexo masculino. Não há anomalias ou tendências preocupantes nesse aspecto, e os dados refletem uma distribuição consistente entre os sexos.

### **Análise dos Sintomas**

Conforme observado na Tabela A.3, os sintomas mais prevalentes entre os pacientes atendidos são febre, mialgia e cefaleia, com uma taxa de aparição superior a 77%. Esses achados já eram esperados, conforme discutido previamente com os pesquisadores, dado que esses sintomas são marcadores clássicos de infecção por dengue.

Outros sintomas, como dor retroocular, náusea, exantema, vômito, artralgia e dor nas costas, aparecem com uma frequência moderada, variando entre 18% e 7% dos casos. É interessante notar que essa sequência de sintomas reflete um padrão típico em

quadros clínicos de dengue, com sinais menos frequentes à medida que a gravidade ou especificidade dos sintomas aumenta.

Por outro lado, sintomas como artrite, petéquia, leucopenia, laco (teste do laço positivo) e conjuntivite são raramente relatados, aparecendo em menos de 3% dos pacientes. Essa baixa prevalência pode estar relacionada tanto à especificidade desses sintomas quanto à dificuldade de diagnóstico ou registro durante o atendimento, já que eles nem sempre são prontamente observáveis ou reportados pelos pacientes.

### **Análise de Comorbidades**

A análise da variável de comorbidades, conforme Tabela A.4, revela que a condição mais frequentemente relatada é a hipertensão, presente em 2,4% dos casos, seguida pelo diabetes, com 1% dos pacientes. Outras comorbidades aparecem em menos de 0,3% dos casos.

Esse padrão pode sugerir uma possível falha no preenchimento dessa variável, uma vez que comorbidades como diabetes e hipertensão geralmente apresentam prevalência mais alta em populações adultas. A baixa frequência reportada pode indicar que, em muitos atendimentos, essas condições não foram devidamente indagadas ou registradas, levando a um sub-relato. Essa questão pode impactar a qualidade da análise de risco associado a comorbidades e sugere que o preenchimento mais completo ou preciso deve ser incentivado nas futuras coletas de dados.

### **Análise da Evolução**

Quanto à variável de evolução dos casos (conforme Tabela A.5), observa-se que em 99,2% dos casos houve cura. Entretanto, 0,6% dos casos foram registrados como "ignorado", o que, apesar de pequeno, pode impactar a análise, especialmente em cenários de avaliação da gravidade dos casos.

Casos de óbito são raros, com 0,2% reportados como óbito por outras causas. Há, entretanto, 57 casos de óbito em investigação e óbito pelo agravo (dengue), que demandam uma análise mais detalhada. Estes casos serão explorados de maneira mais

minuciosa em etapas posteriores do estudo, a fim de identificar possíveis padrões de comportamento ou características que possam estar associadas a esses desfechos mais graves.

### **Análise da Classificação**

A análise da variável classificação (conforme Tabela A.6) indica que 59,3% dos atendimentos foram classificados como dengue, enquanto 37,7% dos casos foram descartados e 2,1% foram classificados como não identificados.

Dada essa distribuição, um modelo de regressão logística será empregado na próxima etapa da análise para avaliar os riscos relativos e identificar os fatores associados à distinção entre um caso descartado e um caso confirmado de dengue. A compreensão dos fatores de risco e de proteção permitirá uma melhor gestão e triagem de casos suspeitos, contribuindo para decisões mais assertivas no tratamento e controle da doença.

### **Análise de Óbitos**

Por fim, vamos trazer mais detalhes sobre os 57 pacientes que vieram a óbito. Conforme Tabela A.7, 40 (70,2%) destes pacientes foram diagnosticados com dengue grave, sendo que temos na amostra geral 80 pacientes diagnosticados com dengue grave, indicando uma taxa de óbito de 50% para essa classificação.

A partir do histograma da idade desses pacientes na Figura B.3, vemos uma distribuição deslocada mais para pacientes mais velhos em comparação ao histograma da idade da amostra geral na Figura B.1. Outro ponto a ser destacado é que 19 dos 57 óbitos eram pessoas com menos de 40 anos.

Nota-se também, pela Tabela A.8 que a comorbidade mais frequente nos pacientes que vieram a óbito é a hipertensão, com 22 casos e em segundo lugar a diabetes com 13 casos. Notamos também que 33 pacientes possuem algum tipo de comorbidade e que dos 13 casos de diabetes, 10 também são hipertensos.

## 6. Modelo de regressão logística

A seguir, foi realizado o ajuste de um modelo de regressão logística (Paula, 2004), com resposta binomial, cuja variável resposta é a classificação final: dengue ou descartado. A Tabela A.9 apresenta os coeficientes das variáveis explicativas e seus valores-p, nela é possível verificar que todas as variáveis preditoras possuem efeito significativo. Já a Figura B.4 apresenta uma análise de resíduos do modelo para verificar a qualidade do ajuste comparando com o quantil da distribuição Normal, evidenciando que o modelo é satisfatório.

A partir do modelo ajustado, é possível obter as estimativas das razões de chances das variáveis. Com isso, as estimativas das razões de chances, seus intervalos de confiança e interpretações encontram-se abaixo:

- Febre: a presença de febre aumenta cerca de 10% (IC = 6% a 14%) a chance de diagnóstico de dengue, em relação a quem não apresentou febre, considerando as outras variáveis constantes.
- Mialgia: a presença de mialgia aumenta cerca de 20% (IC: 15% a 24%) a chance de diagnóstico de dengue, em relação a quem não apresentou esse sintoma, considerando as demais variáveis constantes.
- Cefaleia: a presença de cefaleia aumenta cerca de 21% (IC: 17% a 25%) a chance de diagnóstico de dengue, em relação a quem não apresentou cefaleia, considerando as demais variáveis constantes.
- Exantema: a presença de exantema aumenta cerca de 53% (IC: 47% a 59%) a chance de diagnóstico de dengue, em relação a quem não apresentou exantema, considerando as demais variáveis constantes.
- Dor nas costas: a presença de dor nas costas aumenta cerca de 18% (IC: 12% a 23%) a chance de diagnóstico de dengue, em relação a quem não apresentou esse sintoma, considerando as demais variáveis constantes.

- Petéquias: a presença de petéquias aumenta cerca de 20% (IC: 11% a 30%) a chance de diagnóstico de dengue, em relação a quem não apresentou petéquias, considerando as demais variáveis constantes.
- Leucopenia: a presença de leucopenia aumenta cerca de 21% (IC: 12% a 30%) a chance de diagnóstico de dengue, em relação a quem não apresentou leucopenia, considerando as demais variáveis constantes
- Laço: a realização da prova do laço aumenta em aproximadamente 71% (IC = [54%; 88%]) a chance de diagnóstico de dengue, em comparação quem não foi diagnosticado com dengue, considerando as outras variáveis constantes.
- Dor retroorbital: a presença de dor retroorbital aumenta em aproximadamente 55% (IC = [50%; 61%]) a chance de diagnóstico de dengue, em comparação quem não foi diagnosticado com dengue, considerando as outras variáveis constantes.

## 7. Análise de séries temporais

Os pesquisadores desejam avaliar o comportamento dos dados de atendimento e de notificação ao longo do tempo, a fim de verificar a relação entre elas e realizar previsões para as contagens de notificações. Nesse estudo será realizada uma análise de séries temporais (Box e Jenkins, 1994). A unidade de tempo a ser considerada é a semana epidemiológica (calendário SINAN). Vale ressaltar, neste momento, que os dados de atendimento datam de Janeiro/2022 a Maio/2024, enquanto os dados de notificação datam de Janeiro/2008 a Agosto/2024. Conforme a Figura B.5, é possível verificar que a série de notificações não apresenta média e variância constantes, havendo momentos com alta volatilidade e outros em que a contagem de notificações é estável. Além disso, o gráfico sugere uma possível presença de sazonalidade. As medidas descritivas da série de notificações se encontram na Tabela A.10.

Pela Figura B.6, nota-se que o comportamento da série de atendimento é semelhante, isto é, a série não apresenta indícios de média constante, também havendo períodos em

que a contagem de casos foi significativamente maior. A Tabela A.11 apresenta as medidas descritivas desta série. A Figura B.7 apresenta as duas séries conjuntamente, mas apenas a partir de 2022. Nota-se que as curvas apresentam um comportamento bastante semelhante, evidenciando a forte relação entre as contagens de notificações e de atendimentos.

Uma forma de avaliar a relação entre as duas séries enunciadas é utilizando a técnica de análise de correlação cruzada (Box, e Jenkins, 1994). Conforme Figura B.8, a função de correlação cruzada atinge o seu máximo no lag 0, o que significa que as duas séries têm uma forte correlação no mesmo período, indicando que se movem juntas sem defasagem temporal significativa. Esse resultado sugere uma forte relação contemporânea entre as séries, ou seja, quando uma varia, a outra tende a variar de maneira similar no mesmo instante. Diante disso, para realizar a previsão da série de notificações, é possível utilizar a própria série de notificações como variável preditiva, uma vez que ela já captura de maneira eficiente o comportamento temporal necessário para modelagem, sem a necessidade de recorrer à série de atendimentos.

## **7.1 Modelagem de séries temporais**

A modelagem da série de notificações foi realizada calculando-se o número mensal de notificações como mostrado em vários trabalhos em Gabriel et al. (2019). Conforme verificado anteriormente, a série de notificações apresenta grande variação e, por isso, realizou-se a transformação logarítmica da mesma para obtenção de uma nova série com média constante. A Figura B.9 apresenta a série de notificações com essa transformação. Pode-se verificar que a série se encontra significativamente mais estável do que a original.

Outro efeito a ser considerado é a possível presença de sazonalidade, conforme citado anteriormente. A Figura B.10 apresenta as curvas de notificações de 2008 a 2024. Fica evidenciado que as curvas não apresentam o mesmo comportamento anualmente, pois houve diversos anos em que a quantidade de notificações ao longo das semanas

epidemiológicas se manteve estável. Ao mesmo tempo, nota-se que os picos das curvas de notificações acontecem entre as semanas 1 e 22.

Sendo assim, realizou-se a seguir a diferenciação da curva log-notificações em  $d = 12$ , para verificação do efeito de sazonalidade anual, ou seja, calculamos o valor da série no tempo  $t$  menos o valor da série 12 meses antes. Conforme Figura B.11, fica evidente o efeito de sazonalidade na série de notificações, que deverá ser considerado no modelo ajustado. Em seguida, foi realizado o ajuste de um modelo SARIMA (1,0,1) (2,1,2) [12], cujos coeficientes e respectivos erros padrão se encontram na Tabela A.12. O modelo em questão possuem as seguintes componentes:

- ARIMA(1,0,1):
  - Um componente autorregressivo de ordem 1, com coeficiente  $ar1 = 0,826$ , com erro padrão  $ep = 0,045$ , o que significa que o coeficiente é significativo.
  - Um componente de média móvel de ordem 1, com coeficiente  $ma1 = 0,483$ , e erro padrão  $ep = 0,066$ , mostrando que o coeficiente é significativo.
- Sazonalidade - SARIMA(2,1,2)[12]:
  - Dois componentes autorregressivos sazonais, com coeficientes  $sar1 = -0,735$ , erro padrão  $ep1 = 0,150$ ; e  $sar2 = -0,33$ , com  $ep2 = 0,088$ . Assim, ambos os coeficientes são significativos.
  - Dois componentes de médias móveis sazonais, com coeficientes  $sma1 = -0,171$ ,  $ep1 = 0,160$ ; e  $sma2 = -0,558$ ,  $ep2 = 0,169$ . Assim, ambos os coeficientes são significativos.

A Figura B.12 apresenta a série temporal dos resíduos do modelo. É possível verificar que há indícios de variância e média constante zero, semelhante ao comportamento do ruído branco, conforme esperado. Além disso, o valor-p do teste de Ljung-Box para 14 defasagens (Box e Jenkins, 1994) obtido foi de  $p=0,501$ , o que leva à não rejeição da hipótese nula, de que a autocorrelação entre os resíduos até 14 defasagens é zero. A Figura B.13 apresenta o gráfico da função de autocorrelação dos resíduos, confirmando

a informação verificada no teste de Ljung-Box, pois os valores de autocorrelação encontram-se dentro do intervalo de confiança em todos os lags. Por fim, o gráfico QQ plot dos resíduos na Figura B.14 indica uma certa homocedasticidade dos resíduos, indicando boa qualidade do ajuste realizado.

Com indícios suficientes de que o modelo está bem ajustado, segue-se para a realização de previsões em um intervalo de 12 meses, que podem ser verificadas na Figura B.15. No entanto, como o modelo foi ajustado considerando a curva do logaritmo do número de notificações, é necessário aplicar uma função exponencial para obtenção da previsão das contagens de notificações. Os resultados dessa aplicação estão na Figura B.16. Nota-se que as previsões acompanham bem a tendência geral da série histórica, com maior precisão em períodos de baixa notificação. Durante os picos epidêmicos, observa-se maior variabilidade nas previsões, indicando maior incerteza nesse intervalo, o que é comum em situações de maior volatilidade dos dados. Ainda assim, o modelo captura adequadamente os padrões sazonais e permite uma estimativa confiável.

## 8. Conclusões

A análise descritiva destacou a qualidade geral dos dados, evidenciando preenchimento completo em variáveis como sexo e idade, mas revelando desafios significativos, especialmente em sintomas e comorbidades, com altos percentuais de dados faltantes. A melhoria nos registros a partir de 2017, com menos de 5% de dados ausentes, proporcionou uma base mais robusta para análise, totalizando 55 mil registros. Observou-se que febre, mialgia e cefaleia são os sintomas mais prevalentes, enquanto hipertensão e diabetes se destacam como as comorbidades mais relatadas, embora com taxas relativamente baixas, indicando possível subnotificação.

A partir da modelagem por regressão logística, foi possível identificar fatores associados ao diagnóstico de dengue. Febre, mialgia e cefaleia aumentam significativamente as chances de confirmação do diagnóstico, assim como sintomas específicos como exantema, dor retroorbital e a realização da prova do laço. O modelo apresentou bom ajuste, evidenciado pela análise de resíduos. Essa abordagem quantitativa complementa a análise descritiva, permitindo uma compreensão mais detalhada dos fatores que diferenciam casos confirmados de descartados, contribuindo para estratégias mais eficazes de controle da dengue.

Com relação às séries temporais, conclui-se que a defasagem entre as curvas de notificações e de atendimentos é muito próxima de zero, de forma que a previsão da contagem de notificações pode ser realizada considerando ela mesma. Além disso, a aplicação de transformação em log provou-se útil para obtenção de estacionaridade e facilitar a modelagem. Também, a série de notificações apresenta fortes indícios de sazonalidade, que deve ser considerada no modelo estatístico a ser ajustado. Por fim, o modelo SARIMA ajustado apresentou um bom desempenho ao ser feita a análise de resíduos, embora as previsões possam apresentar maior variabilidade durante os picos de notificação. Ainda assim, o modelo foi capaz de capturar os padrões sazonais e fornecer estimativas consistentes, especialmente em períodos de baixa notificação, podendo ser uma ferramenta eficaz para o monitoramento e planejamento de ações relacionadas à dengue.

# **APÊNDICE A**

## **Tabelas**

**Tabela A.1** Percentual de dados faltantes por variável

| Variável         | Valores Faltantes | Percentual Faltantes |
|------------------|-------------------|----------------------|
| idade            | 0                 | 0,0%                 |
| sexo             | 0                 | 0,0%                 |
| data notificação | 0                 | 0,0%                 |
| classificação    | 192               | 0,1%                 |
| evolução         | 50.683            | 17,6%                |
| hospitalização   | 87.169            | 30,3%                |
| febre            | 135.913           | 47,3%                |
| mialgia          | 135.913           | 47,3%                |
| cefaleia         | 135.913           | 47,3%                |
| exantema         | 135.913           | 47,3%                |
| vômito           | 135.913           | 47,3%                |
| náusea           | 135.913           | 47,3%                |
| dor costas       | 135.913           | 47,3%                |
| conjuntivite     | 135.913           | 47,3%                |
| artrite          | 135.913           | 47,3%                |
| artralgia        | 135.913           | 47,3%                |
| petequia         | 135.913           | 47,3%                |
| leucopenia       | 135.913           | 47,3%                |
| laco             | 135.913           | 47,3%                |
| dor retro        | 135.913           | 47,3%                |
| diabetes         | 135.913           | 47,3%                |
| hematolog        | 135.913           | 47,3%                |
| hepatopat        | 135.913           | 47,3%                |
| renal            | 135.913           | 47,3%                |
| hipertensa       | 135.913           | 47,3%                |
| auto imune       | 135.913           | 47,3%                |
| sorotipo         | 285.831           | 99,5%                |

**Tabela A.2** Percentual de dados faltantes ao longo dos anos

| Variavel       | 2008 | 2009 | 2010 | 2011 | 2012 | 2013 | 2014 | 2015 | 2016 | 2017 | 2018 | 2019 | 2020 | 2021 | 2022 | 2023 | 2024 |
|----------------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|
| febre          | 100% | 100% | 100% | 100% | 100% | 100% | 100% | 81%  | 19%  | 1,4% | 0,7% | 3,0% | 2,3% | 0,8% | 0,1% | 0,0% | 0,0% |
| mialgia        | 100% | 100% | 100% | 100% | 100% | 100% | 100% | 81%  | 19%  | 1,4% | 0,7% | 3,0% | 2,3% | 0,8% | 0,1% | 0,0% | 0,0% |
| cefaleia       | 100% | 100% | 100% | 100% | 100% | 100% | 100% | 81%  | 19%  | 1,4% | 0,7% | 3,0% | 2,3% | 0,8% | 0,1% | 0,0% | 0,0% |
| exantema       | 100% | 100% | 100% | 100% | 100% | 100% | 100% | 81%  | 19%  | 1,4% | 0,7% | 3,0% | 2,3% | 0,8% | 0,1% | 0,0% | 0,0% |
| vomito         | 100% | 100% | 100% | 100% | 100% | 100% | 100% | 81%  | 19%  | 1,4% | 0,7% | 3,0% | 2,3% | 0,8% | 0,1% | 0,0% | 0,0% |
| nausea         | 100% | 100% | 100% | 100% | 100% | 100% | 100% | 81%  | 19%  | 1,4% | 0,7% | 3,0% | 2,3% | 0,8% | 0,1% | 0,0% | 0,0% |
| dor costas     | 100% | 100% | 100% | 100% | 100% | 100% | 100% | 81%  | 19%  | 1,4% | 0,7% | 3,0% | 2,3% | 0,8% | 0,1% | 0,0% | 0,0% |
| conjuntivite   | 100% | 100% | 100% | 100% | 100% | 100% | 100% | 81%  | 19%  | 1,4% | 0,7% | 3,0% | 2,3% | 0,8% | 0,1% | 0,0% | 0,0% |
| artrite        | 100% | 100% | 100% | 100% | 100% | 100% | 100% | 81%  | 19%  | 1,4% | 0,7% | 3,0% | 2,3% | 0,8% | 0,1% | 0,0% | 0,0% |
| artralgia      | 100% | 100% | 100% | 100% | 100% | 100% | 100% | 81%  | 19%  | 1,4% | 0,7% | 3,0% | 2,3% | 0,8% | 0,1% | 0,0% | 0,0% |
| petequia       | 100% | 100% | 100% | 100% | 100% | 100% | 100% | 81%  | 19%  | 1,4% | 0,7% | 3,0% | 2,3% | 0,8% | 0,1% | 0,0% | 0,0% |
| leucopenia     | 100% | 100% | 100% | 100% | 100% | 100% | 100% | 81%  | 19%  | 1,4% | 0,7% | 3,0% | 2,3% | 0,8% | 0,1% | 0,0% | 0,0% |
| laco           | 100% | 100% | 100% | 100% | 100% | 100% | 100% | 81%  | 19%  | 1,4% | 0,7% | 3,0% | 2,3% | 0,8% | 0,1% | 0,0% | 0,0% |
| dor retro      | 100% | 100% | 100% | 100% | 100% | 100% | 100% | 81%  | 19%  | 1,4% | 0,7% | 3,0% | 2,3% | 0,8% | 0,1% | 0,0% | 0,0% |
| diabetes       | 100% | 100% | 100% | 100% | 100% | 100% | 100% | 81%  | 19%  | 1,4% | 0,7% | 3,0% | 2,3% | 0,8% | 0,1% | 0,0% | 0,0% |
| hematolog      | 100% | 100% | 100% | 100% | 100% | 100% | 100% | 81%  | 19%  | 1,4% | 0,7% | 3,0% | 2,3% | 0,8% | 0,1% | 0,0% | 0,0% |
| hepatopat      | 100% | 100% | 100% | 100% | 100% | 100% | 100% | 81%  | 19%  | 1,4% | 0,7% | 3,0% | 2,3% | 0,8% | 0,1% | 0,0% | 0,0% |
| renal          | 100% | 100% | 100% | 100% | 100% | 100% | 100% | 81%  | 19%  | 1,4% | 0,7% | 3,0% | 2,3% | 0,8% | 0,1% | 0,0% | 0,0% |
| hipertensa     | 100% | 100% | 100% | 100% | 100% | 100% | 100% | 81%  | 19%  | 1,4% | 0,7% | 3,0% | 2,3% | 0,8% | 0,1% | 0,0% | 0,0% |
| auto imune     | 100% | 100% | 100% | 100% | 100% | 100% | 100% | 81%  | 19%  | 1,4% | 0,7% | 3,0% | 2,3% | 0,8% | 0,1% | 0,0% | 0,0% |
| evolução       | 88%  | 49%  | 12%  | 18%  | 21%  | 59%  | 4%   | 12%  | 18%  | 3,3% | 3,0% | 4,9% | 3,9% | 1,6% | 1,1% | 0,4% | 0,3% |
| hospitalização | 100% | 97%  | 63%  | 28%  | 33%  | 59%  | 26%  | 35%  | 24%  | 37%  | 1,7% | 7,7% | 7,7% | 1,3% | 1,5% | 0,4% | 0,9% |
| sorotipo       | 100% | 100% | 100% | 100% | 100% | 100% | 100% | 99%  | 100% | 100% | 99%  | 100% | 98%  | 100% | 98%  | 99%  | 99%  |

**Tabela A.3** Percentual de ocorrência dos sintomas

| Sintoma      | Contagem Casos | Percentual Aparição |
|--------------|----------------|---------------------|
| febre        | 127.824        | 84,0%               |
| mialgia      | 126.158        | 82,2%               |
| cefaleia     | 120.380        | 77,5%               |
| dor retro    | 27.029         | 17,5%               |
| náusea       | 22.224         | 14,4%               |
| exantema     | 21.093         | 13,3%               |
| artralgia    | 18.210         | 11,3%               |
| vômito       | 16.357         | 10,3%               |
| dor costas   | 10.180         | 6,5%                |
| artrite      | 4.283          | 2,8%                |
| petequia     | 4.021          | 2,7%                |
| leucopenia   | 3.975          | 2,7%                |
| laco         | 3.430          | 1,8%                |
| conjuntivite | 1.833          | 0,9%                |

**Tabela A.4** Percentual de ocorrência das comorbidades

| Comorbidade | Contagem Casos | Percentual Aparicao |
|-------------|----------------|---------------------|
| hipertensa  | 2862           | 2,4%                |
| diabetes    | 1163           | 1,0%                |
| autoimmune  | 317            | 0,3%                |
| hematolog   | 257            | 0,2%                |
| hepatopat   | 253            | 0,2%                |
| renal       | 231            | 0,2%                |

**Tabela A.5** Distribuição da variável evolução por categoria

| Evolução            | Contagem Casos | Percentual |
|---------------------|----------------|------------|
| Cura                | 118.218        | 99,2%      |
| Ignorado            | 695            | 0,6%       |
| Óbito Outras Causas | 233            | 0,2%       |
| Óbito Investigação  | 7              | 0,0%       |
| Óbito pelo Agravado | 50             | 0,0%       |

**Tabela A.6** Distribuição da variável classificação por categoria

| Classificação   | Contagem Casos | Percentual |
|-----------------|----------------|------------|
| Dengue          | 72.733         | 59,3%      |
| Descartado      | 46289          | 37,7%      |
| Não dentificado | 2518           | 2,1%       |
| Dengue Alarme   | 1087           | 0,9%       |
| Dengue Grave    | 80             | 0,1%       |

**Tabela A.7** Distribuição da variável classificação por categoria para óbitos

| Classificação    | Contagem Casos | Percentual |
|------------------|----------------|------------|
| Dengue           | 40             | 70,2%      |
| Descartado       | 9              | 15,8%      |
| Não identificado | 5              | 8,8%       |
| Dengue Alarme    | 2              | 3,5%       |
| Dengue Grave     | 1              | 1,8%       |

**Tabela A.8** Percentual de ocorrência das comorbidades para óbitos

| Comorbidade | Contagem Casos | Percentual Aparicao |
|-------------|----------------|---------------------|
| hipertensa  | 22             | 38,6%               |
| diabetes    | 13             | 22,8%               |
| renal       | 4              | 7,0%                |
| hematolog   | 3              | 5,3%                |
| hepatopat   | 3              | 5,3%                |
| autoimune   | 3              | 5,3%                |

**Tabela A.9** Coeficientes do modelo de regressão logística com classificação final como variável resposta

| Coeficiente | Estimativa | Valor-p |
|-------------|------------|---------|
| Intercepto  | -0,09      | <0,01   |
| Febre       | 0,09       | <0,01   |
| Mialgia     | 0,18       | <0,01   |

|                  |      |       |
|------------------|------|-------|
| Cefaleia         | 0,19 | <0,01 |
| Exantema         | 0,42 | <0,01 |
| Dor nas Costas   | 0,16 | <0,01 |
| Petéquias        | 0,19 | <0,01 |
| Leucopenia       | 0,19 | <0,01 |
| Prova do Laço    | 0,53 | <0,01 |
| Dor retroorbital | 0,44 | <0,01 |

**Tabela A.10** Medidas descritivas da série de notificações

| Mínimo | 1º Quartil | Mediana | Média | 3º Quartil | Máximo | Variância |
|--------|------------|---------|-------|------------|--------|-----------|
| 1      | 51         | 99      | 330   | 225        | 5.914  | 449.859   |

**Tabela A.11** Medidas descritivas da série de atendimentos

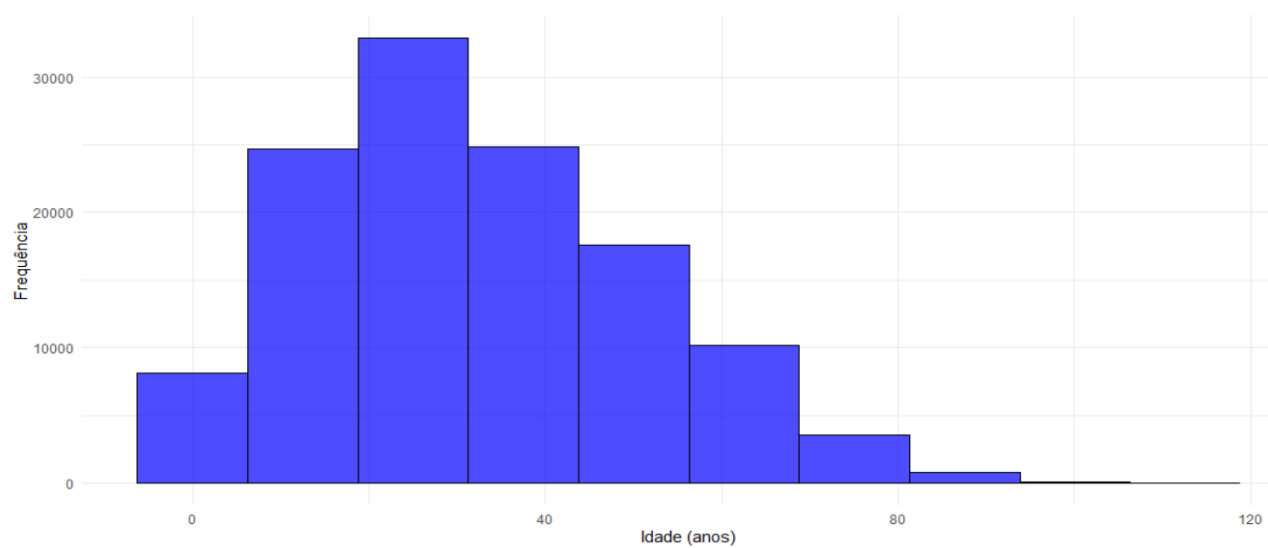
| Mínimo | 1º Quartil | Mediana | Média | 3º Quartil | Máximo | Variância |
|--------|------------|---------|-------|------------|--------|-----------|
| 63     | 142        | 229     | 334   | 475        | 1.178  | 58.536    |

**Tabela A.12** Coeficientes e erros padrão do modelo SARIMA

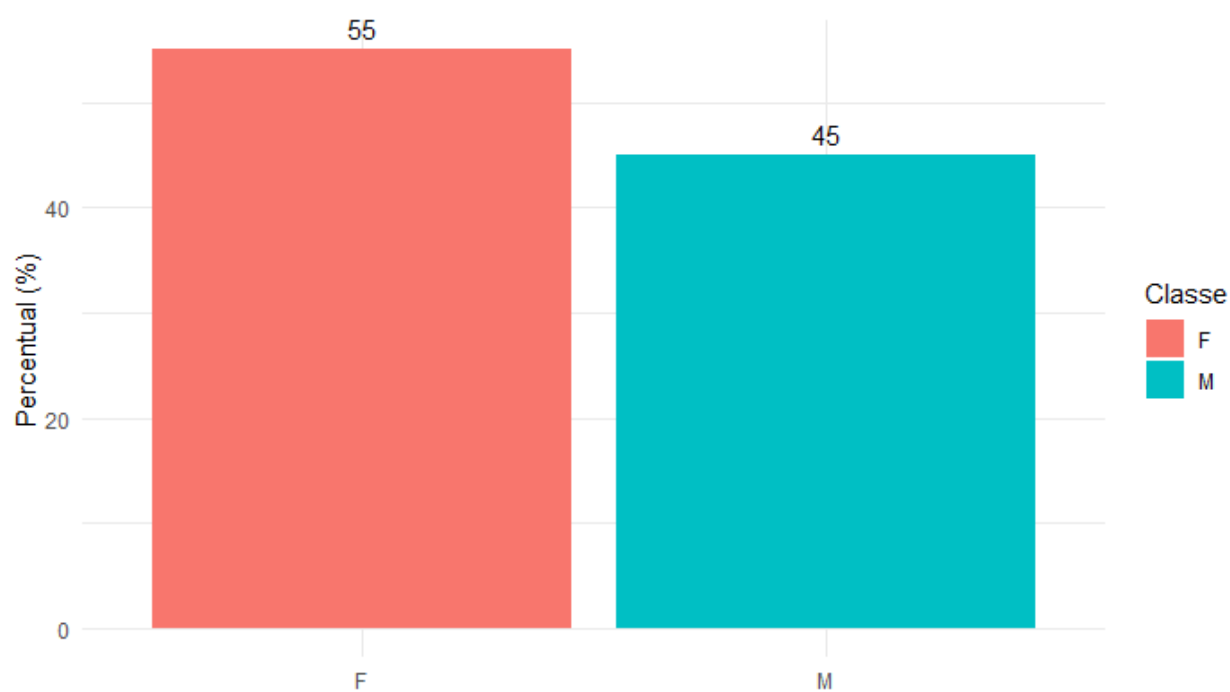
| Coeficiente | Valor estimado | Erro padrão |
|-------------|----------------|-------------|
| AR(1)       | 0,83           | 0,05        |
| MA(1)       | 0,48           | 0,07        |
| SAR(1)      | -0,73          | 0,15        |
| SAR(2)      | -0,32          | 0,09        |
| SMA(1)      | -0,17          | 0,16        |
| SMA(2)      | -0,56          | 0,17        |

# **APÊNDICE B**

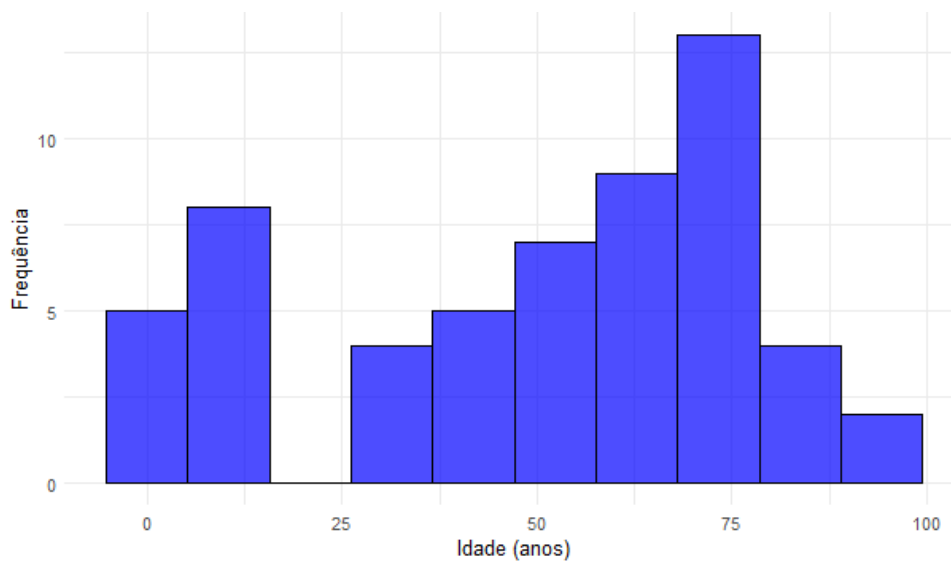
## **Figuras**



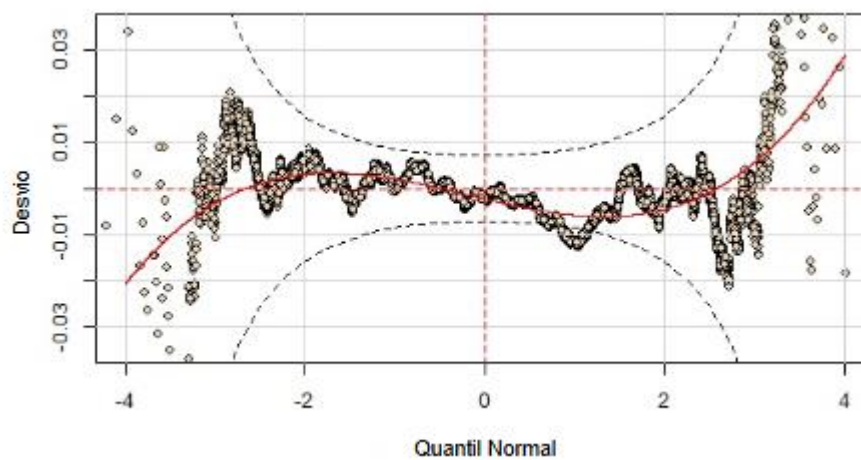
**Figura B.1** Histograma da idade



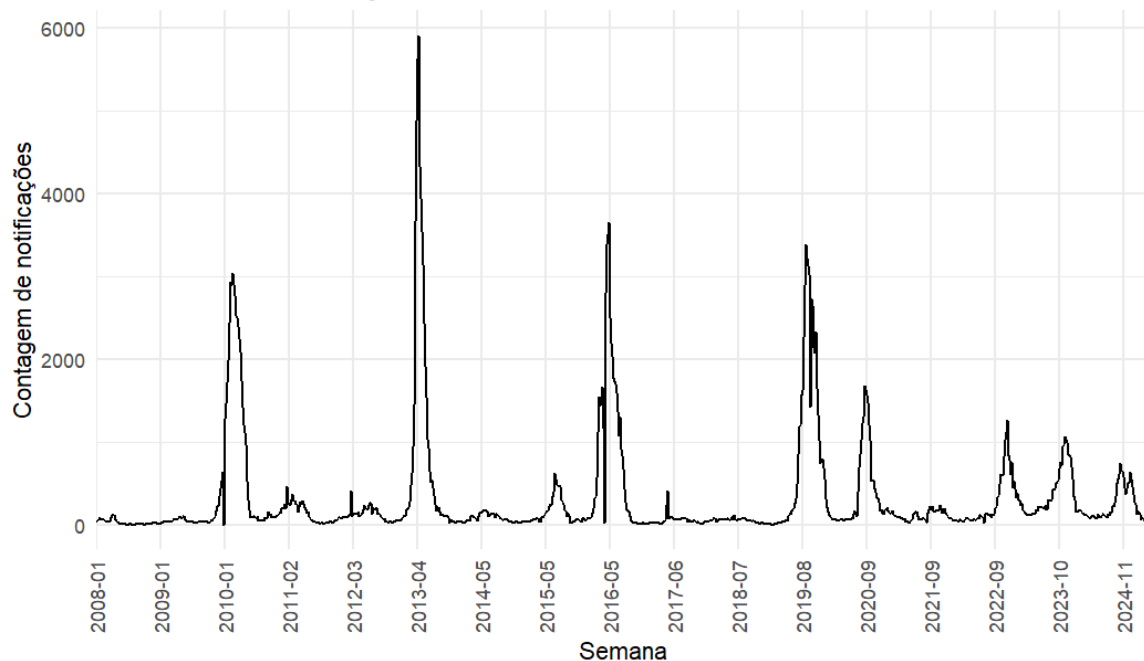
**Figura B.2** Distribuição da variável sexo dos pacientes



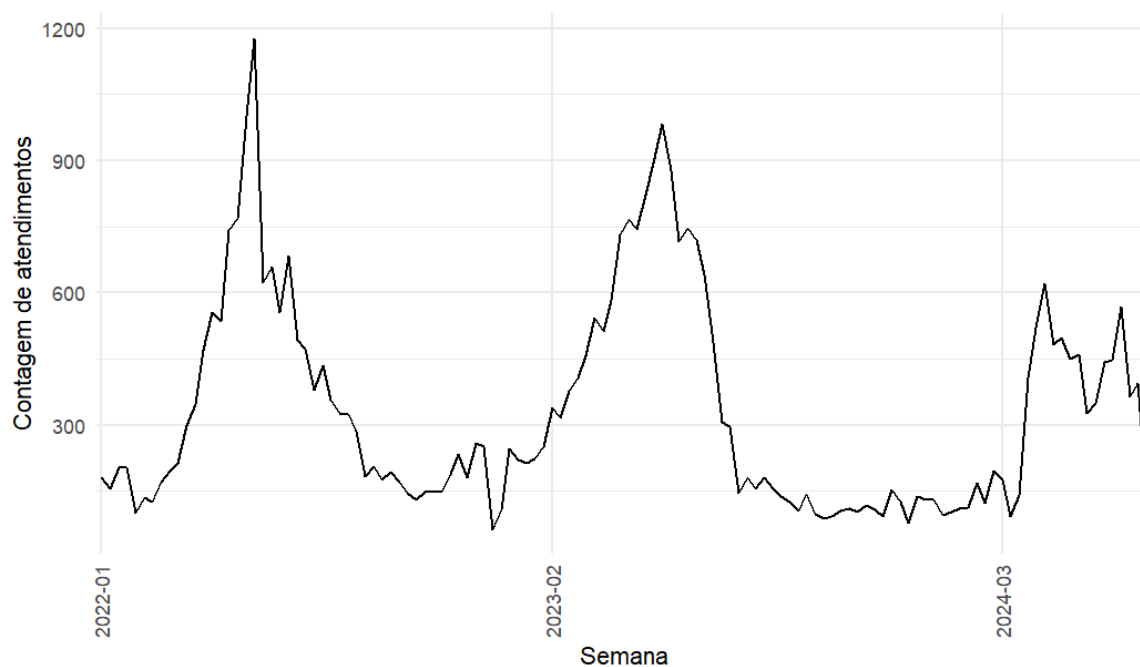
**Figura B.3** Histograma da idade para óbitos



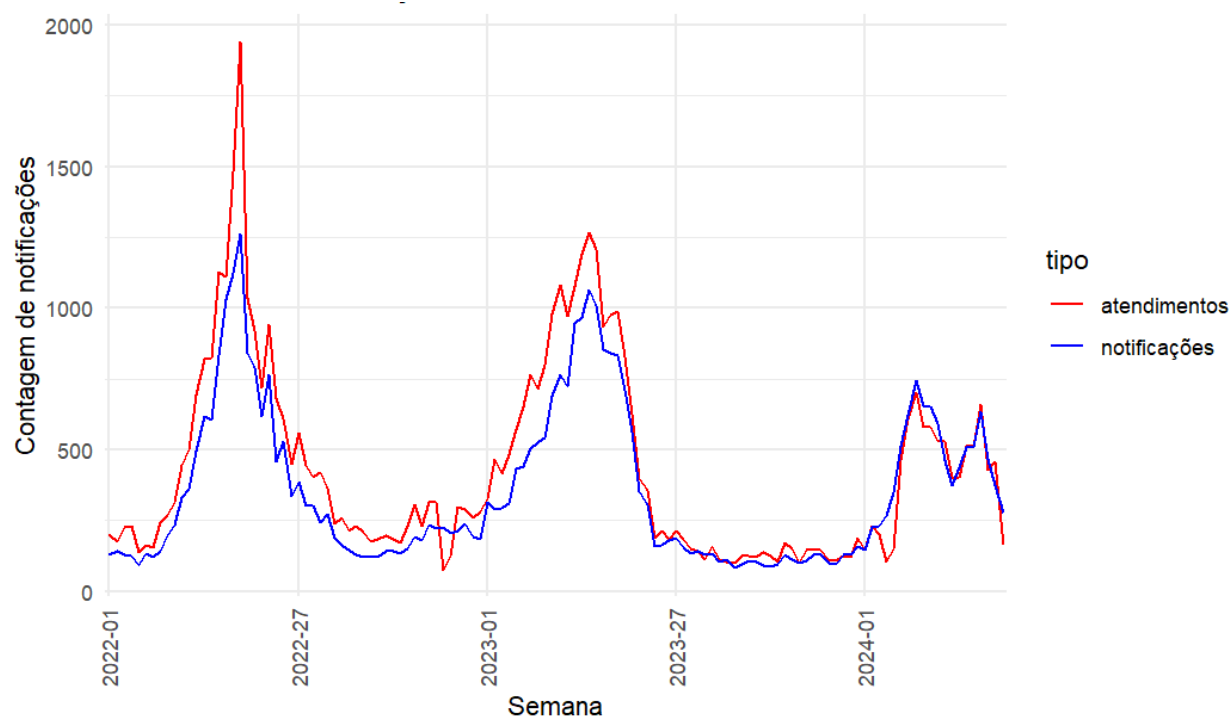
**Figura B.4** Worm Plot dos resíduos do modelo de regressão logística



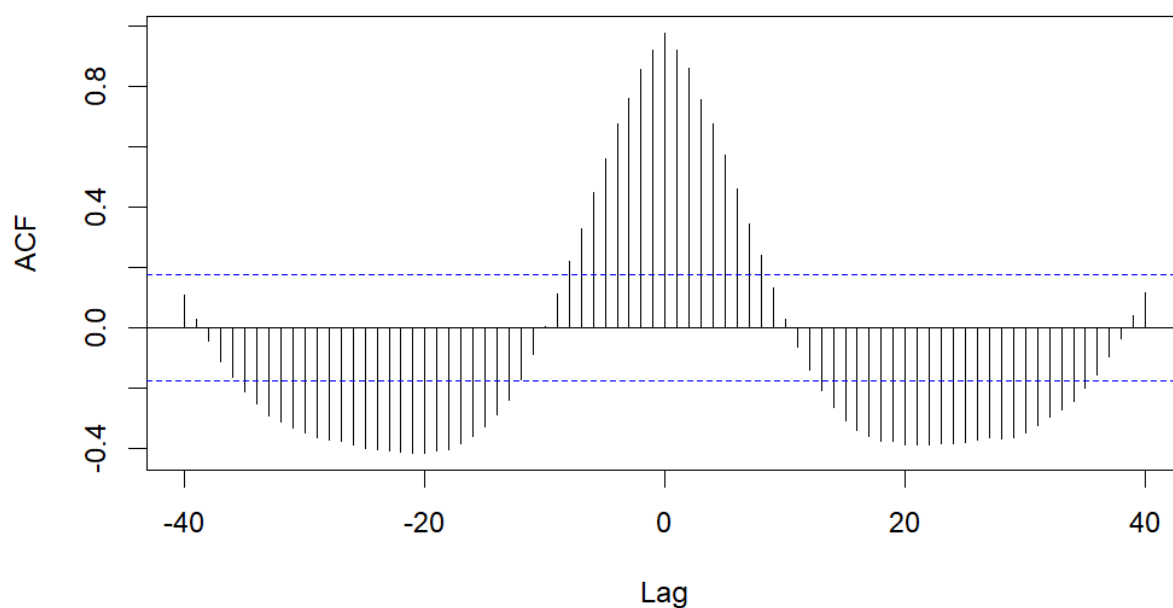
**Figura B.5** Curva de contagem de notificações por semana epidemiológica



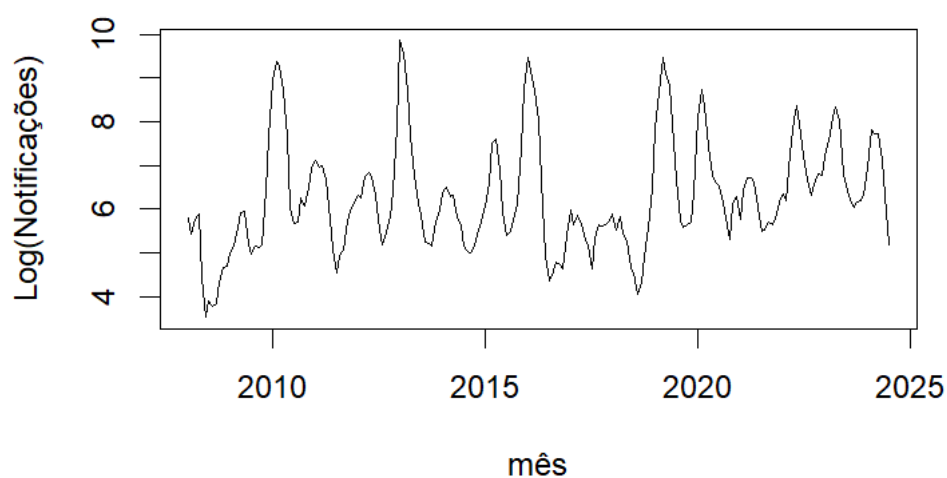
**Figura B.6** Curva de contagem de atendimentos por semana epidemiológica



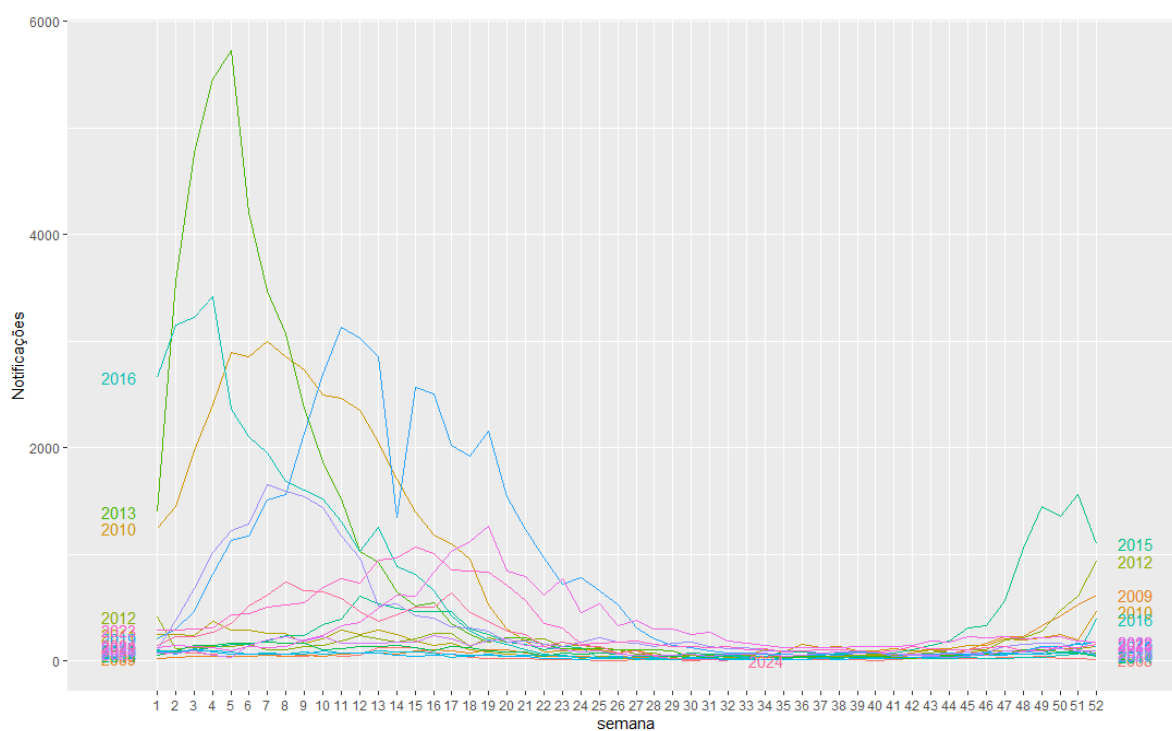
**Figura B.7** Curvas de contagem de atendimento e de notificações por semana epidemiológica



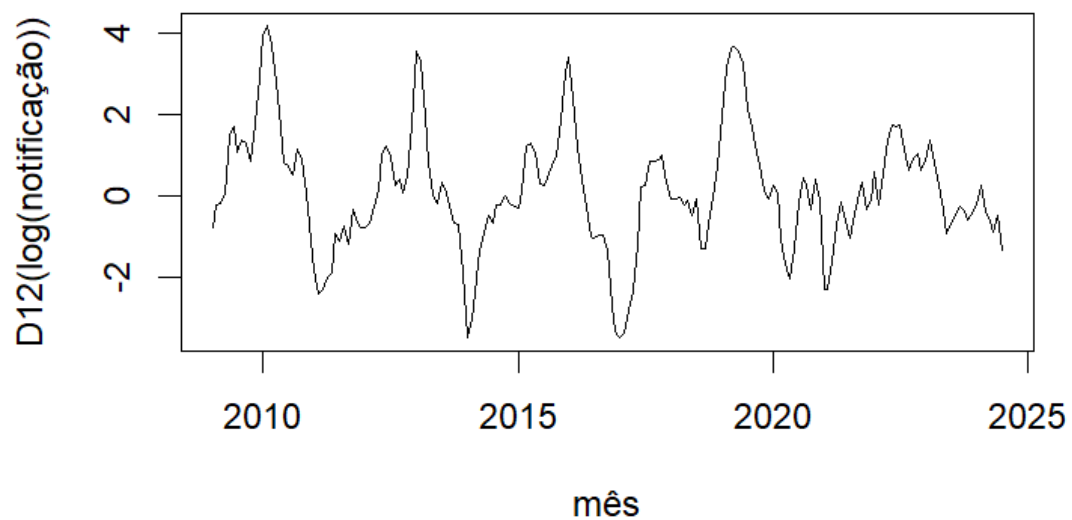
**Figura B.8** Gráfico da função de correlação cruzada das séries de notificações e de atendimento



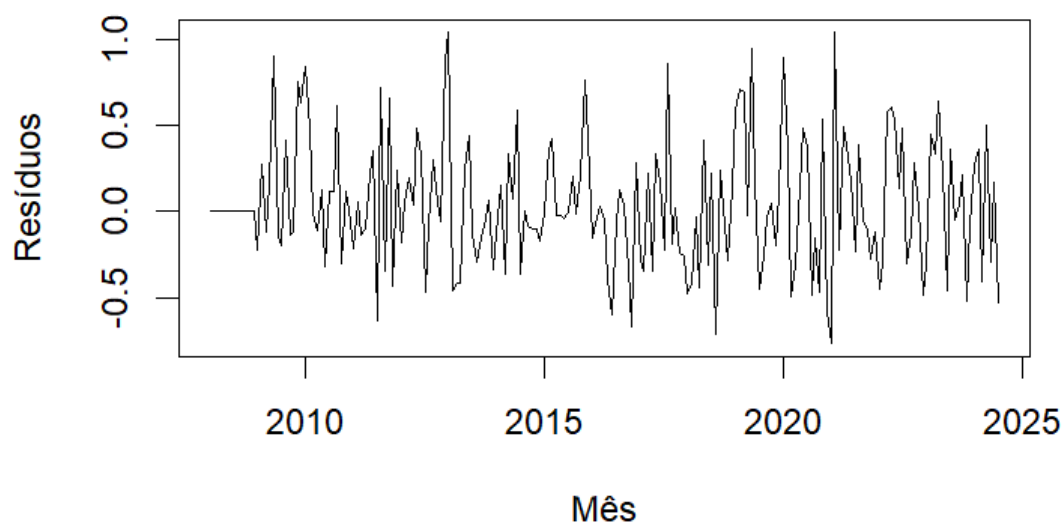
**Figura B.9** Curva de log-notificações



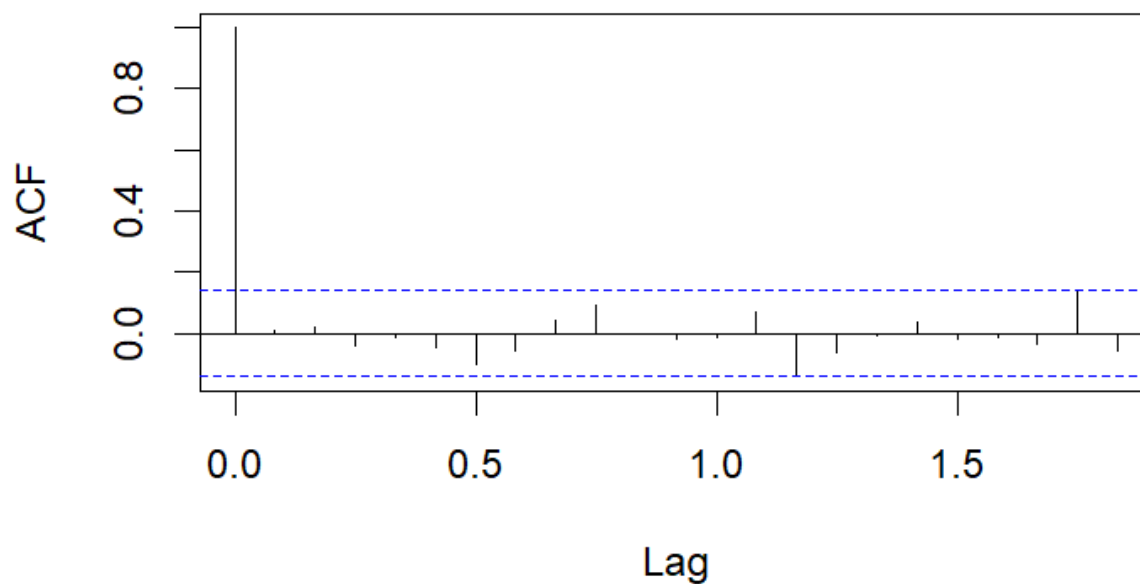
**Figura B.10** Curvas de contagem notificações por semana epidemiológica, de 2008 a 2024, visão sazonal



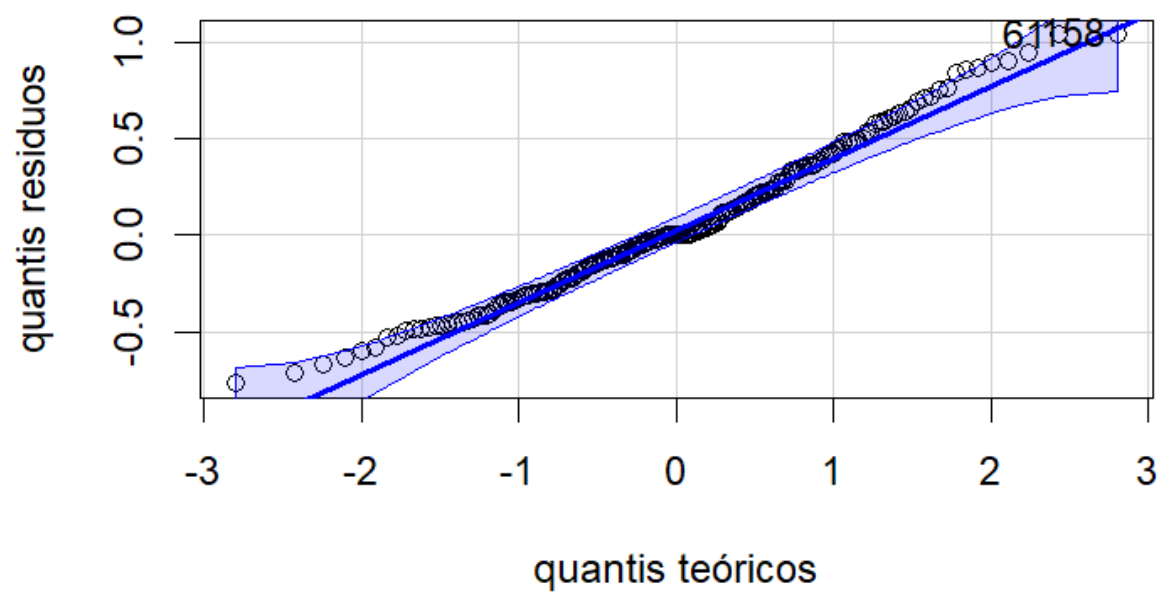
**Figura B.11** Curva de log-notificações diferenciada em  $d = 12$



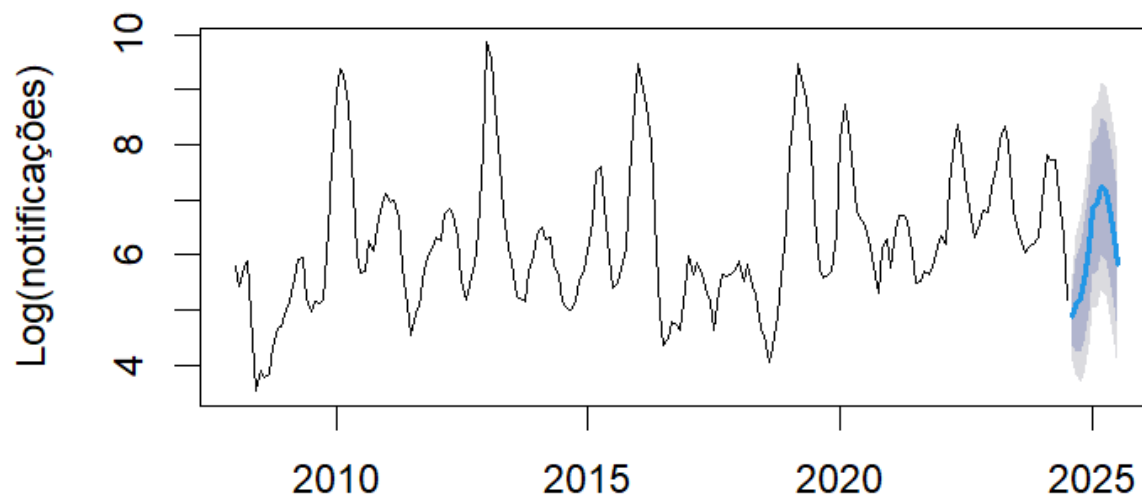
**Figura B.12** Série temporal dos resíduos do modelo SARIMA ajustado



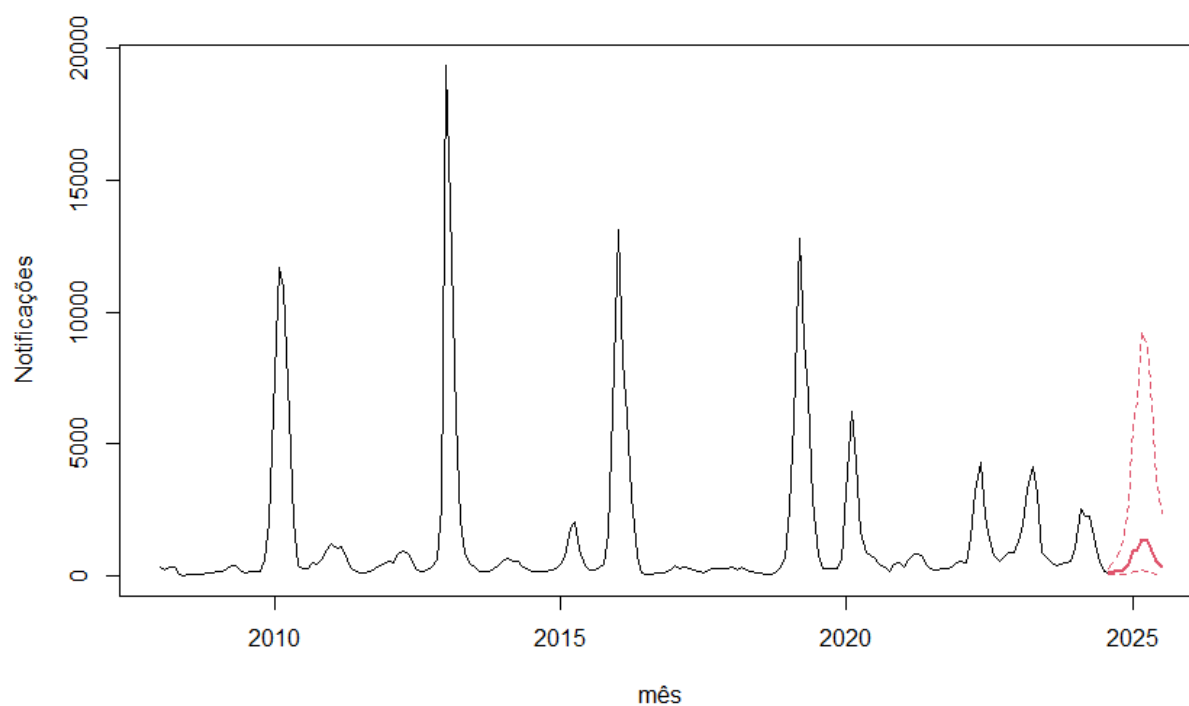
**Figura B.13** Gráfico da função de autocorrelação dos resíduos do modelo SARIMA



**Figura B.14** Gráfico QQ plot dos resíduos do modelo SARIMA



**Figura B.15** Curva de log-notificações com previsão de 12 meses



**Figura B.16** Curva de notificações com previsão de 12 meses