

UNIVERSIDADE FEDERAL DE MINAS GERAIS  
INSTITUTO DE CIÊNCIAS EXATAS  
DEPARTAMENTO DE ESTATÍSTICA  
ESTATÍSTICA

DANIELLY SANTOS SEVERINO

**PREVISÃO DE SÉRIES EPIDEMIOLÓGICAS INCORPORANDO  
ATRASO NA NOTIFICAÇÃO**

BELO HORIZONTE  
2023

DANIELLY SANTOS SEVERINO

**PREVISÃO DE SÉRIES EPIDEMIOLÓGICAS INCORPORANDO  
ATRASO NA NOTIFICAÇÃO**

Monografia apresentada ao Departamento de Estatística da Universidade Federal de Minas Gerais (UFMG), como parte dos requisitos necessários à obtenção do título de bacharel em Estatística.

Orientador: Dani Gamerman

Coorientadora: Izabel Nolau

BELO HORIZONTE

2023

## **Agradecimentos**

Algumas épocas da vida a gente trabalha e estuda tanto que o tempo com a família é raro, quando estamos em casa só queremos aproveitar. O engraçado e mais contraditório é que é tudo por eles. Não reclamo porque eu vivo um sonho. Todos os dias posso acessar minha fonte de felicidade, quero ser forte como a Tia Lú (Vera Lúcia) em tempos difíceis.

Agradeço por poder olhar nos olhos do meu irmão Lucas e entender melhor a vida, aprender a me dedicar mais. Sou grata pela minha família e amigos que incentivam minha jornada. Me sinto mais livre sempre que me aproximo mais das minhas raízes. Deus, obrigada por ter me tornado a continuação de um sonho do meu pai e da minha mãe, de todos que vieram antes de mim.

Aos meu orientador Dani Gamerman e minha coorientadora Izabel Nolau, agradeço pelos esforços, dedicação, inspiração, apoio e paciência em me auxiliar a concretizar esse trabalho. Agradeço também aos professores que contribuíram para a minha formação, especialmente Cristina Eterovik e Helinton A. Lopes. Por fim, agradeço a Fundação Mendes Pimental (FUMP) por me proporcionar condições financeiras para cursar o início da graduação.

## Resumo

A gestão do risco de doenças infecciosas envolve a notificação e coleta de dados para monitorar a disseminação e identificar tendências das doenças. No entanto, os casos notificados muitas vezes sofrem atraso na notificação, ou seja, os casos são notificados após a incidência da doença. Os atrasos distorcem a relação entre a incidência relatada e a verdadeira incidência da doença. Isso pode ter consequências significativas para a saúde pública e para a eficácia de medidas de medidas de controle. Sem informações adequadas da gravidade da situação atual em relação à doença, as autoridades têm dificuldade em agir de maneira eficaz para identificar e responder adequadamente aos surtos. Sendo assim, é crucial corrigir a informação disponível o mais rápido possível para auxiliar a tomada de decisões como emissões de aviso às autoridades locais e ao público-alvo.

Os modelos hierárquicos Bayesianos são utilizados para lidar com a incerteza em vários níveis dos dados, permitindo combinar informações de diferentes níveis para obter estimativas e previsões mais precisas e robustas. Ao lidar com doenças infecciosas é importante que os modelos sejam capazes de gerar estimativas que recuperem a informação existente com precisão e, especialmente, construir previsões robustas. As previsões podem ser feitas em tempo real (*nowcasting*) com base nos dados disponíveis ou para períodos futuros (*forecasting*) com base no histórico.

Neste trabalho, foram propostos modelos hierárquicos Bayesianos com duas estruturas diferentes para lidar com dados de contagem de incidência de doenças infecciosas e corrigir o atraso nas notificações. Esses modelos são uma forma flexível de corrigir os atrasos nas notificações e permitem gerar previsões e quantificar a incerteza associada. Os modelos propostos foram ilustrados usando dados semanais de incidência de dengue no Rio de Janeiro registrados entre janeiro de 2011 e abril de 2012. Além disso, a técnica de Monte Carlo via Cadeias de Markov (MCMC) foi utilizada para a extração de amostras para aproximar as densidades a posteriori das quantidades de interesse.

**Palavras-chaves:** Estatística; inferência Bayesiana; séries temporais; dados de contagem; atraso na notificação; *Nowcasting*; *Forecasting*; dengue.

## Sumário

## **Introdução**

O sistema de vigilância de doenças infecciosas possui algumas abordagens comuns que muitos países adotam (M. Steven; R. Elliott, 2000). Na maioria dos países a notificação é obrigatória, exigindo que a incidência de doenças infecciosas como tuberculose, hepatite, HIV/AIDS e sarampo seja informada às autoridades de saúde. Os profissionais de saúde, hospitais e laboratórios são responsáveis por relatar os casos suspeitos ou confirmados. Portanto, existe um sistema de notificação e coleta de dados que se caracteriza por estabelecer informações essenciais para monitorar a disseminação e identificar tendências das doenças. Cabe destacar que os sistemas de vigilância epidemiológica ajudam a identificar surtos, avaliar a eficácia das medidas de controle e orientar políticas públicas de saúde. Além disso, a pandemia de COVID-19 trouxe à tona a importância da vigilância de doenças e a necessidade de fortalecer os sistemas de saúde em todo o mundo.

Os casos de doenças infecciosas que são relatados aos profissionais de saúde podem ser analisados através de contagens, sendo que, o número de casos comumente sofre atraso na notificação, ou seja, casos que são relatados após a incidência da doença. Existem diversas situações que podem levar ao atraso na notificação de doenças como tempo até a confirmação do diagnóstico, desafios logísticos, confirmação de exames e barreiras culturais. Dessa forma, os casos notificados distorcem a relação entre a incidência da doença relatada e a verdadeira incidência da doença.

O atraso na notificação de doenças infecciosas pode ter consequências significativas para a saúde pública e para a eficácia de medidas de controle. É extremamente importante corrigir o atraso nas notificações para diminuir o risco de propagação da doença, não retardar o tratamento e cuidados adequados e nem dificultar a detecção de tendências e padrões. Sem informações adequadas da gravidade da situação atual em relação à doença, as autoridades têm dificuldade em agir de forma eficaz para identificar e responder rapidamente aos surtos, permitindo que a infecção entre a população aumente.

Para diversos casos de doenças é comum que os dados sejam potencialmente subnotificados, ou seja, casos de doenças que não foram detectados

ou que foram detectados, mas não notificados. Isso resulta em uma representação imprecisa da verdadeira contagem de doenças. Corrigir a subnotificação de doenças é um desafio complexo que requer uma abordagem abrangente e a implementação de várias estratégias, exigindo fontes adicionais de informação como conhecimento prévio da taxa de subnotificação.

O foco neste trabalho é corrigir o atraso na notificação dos dados. Para isso, foram propostos modelos hierárquicos Bayesianos para dados de contagem de incidência das doenças infecciosas como uma forma flexível de corrigir o atraso nas notificações e quantificar a incerteza associada. A modelagem proposta assume que os dados de contagem seguem uma distribuição de probabilidade Poisson incorporando componentes do efeito do tempo e do efeito das defasagens.

## Estrutura típica dos dados com atraso na notificação

Ao lidar com séries epidemiológicas, existe uma estrutura típica que é utilizada para lidar com dados de contagem da incidência de doenças com problema de atraso na notificação. Através dessa estrutura é possível capturar e analisar adequadamente as informações disponíveis (BASTOS et al, 2019).

A Tabela 1 exibe a estrutura típica de uma base de dados qualquer para contagens com atraso na notificação. As linhas da tabela correspondem ao tempo, enquanto as colunas indicam o período de atraso na notificação.  $T$  é o tempo atual,  $D$  é o atraso máximo relevante e  $H$  é o horizonte máximo de previsão. Os eventos ocorridos no tempo  $t$  com  $d$  unidades de atraso são denotados por  $n_{t,d}$  e, fixado qualquer intervalo de tempo (linha), a quantidade total de eventos é  $N_t = \sum_{d=0}^D n_{t,d}$ .

Os eventos ocorridos entre 1 e  $T - D$  foram integralmente registrados. Já os eventos ocorridos entre  $T - D + 1$  e  $T$  são parcialmente registrados, pois existem eventos que ocorreram e não foram registrados devido ao atraso na notificação. Por fim, as contagens de  $T + 1$  até  $T + H$  são os eventos futuros.

### Estrutura típica dos dados com atraso na notificação

$t \backslash d$	0	1	...	D-1	D	N
1	$n_{1,0}$	$n_{1,1}$		$n_{1,D-1}$	$n_{1,D}$	$N_1$
2	$n_{2,0}$	$n_{2,1}$		$n_{2,D-1}$	$n_{2,D}$	$N_2$
...						
$T - D$	$n_{T-D,0}$	$n_{T-D,1}$		$n_{T-D,D-1}$	$n_{T-D,D}$	$N_{T-D}$
$T - D + 1$	$n_{T-D+1,0}$	$n_{T-D+1,1}$		$n_{T-D+1,D-1}$	$n_{T-D+1,D}$	$N_{T-D+1}$
...						
$T - 1$	$n_{T-1,0}$	$n_{T-1,1}$		$n_{T-1,D-1}$	$n_{T-1,D}$	$N_{T-1}$
$T$	$n_{T,0}$	$n_{T,1}$		$n_{T,D-1}$	$n_{T,D}$	$N_T$
$T + 1$	$n_{T+1,0}$	$n_{T+1,1}$		$n_{T+1,D-1}$	$n_{T+1,D}$	$N_{T+1}$
...						
$T + H$	$n_{T+H,0}$	$n_{T+H,1}$		$n_{T+H,D-1}$	$n_{T+H,D}$	$N_{T+H}$