Caio Fernando Martins Carneiro

Análise comportamental do perfil de consumidores de uma empresa após a utilização da ferramenta Google Ads

Niterói - RJ, Brasil

13 de julho de 2023

Universidade Federal Fluminense

Caio Fernando Martins Carneiro

Análise comportamental do perfil de consumidores de uma empresa após a utilização da ferramenta Google Ads

Trabalho de Conclusão de Curso

Monografia apresentada para obtenção do grau de Bacharel em Estatística pela Universidade Federal Fluminense.

Orientadora: Profa. Dra. Mariana Albi de Oliveira Souza

Niterói - RJ, Brasil

13 de julho de 2023

Universidade Federal Fluminense

Caio Fernando Martins Carneiro

Análise comportamental do perfil de consumidores de uma empresa após a utilização da ferramenta Google Ads

Monografia de Projeto Final de Graduação sob o título *"Análise comportamental do perfil de consumidores de uma empresa após a utilização da ferramenta Google Ads"*, defendida por Caio Fernando Martins Carneiro e aprovada em 13 de julho de 2023, na cidade de Niterói, no Estado do Rio de Janeiro, pela banca examinadora constituída pelos professores:

Profa. Dra. Mariana Albi de Oliveira Souza
Departamento de Estatística – UFF

Profa. Dra. Ana Beatriz Monteiro Fonseca
Departamento de Estatística – UFF

Prof. Dr. Luis Guillermo Coca Velarde
Departamento de Estatística – UFF

Niterói, 13 de julho de 2023

Ficha catalográfica automática - SDC/BIME
Gerada com informações fornecidas pelo autor

C289a Carneiro, Caio Fernando Martins
Análise comportamental do perfil de consumidores de uma empresa após a utilização da ferramenta Google Ads / Caio Fernando Martins Carneiro. - 2023.
47 f.: il.

Orientador: Mariana Albi de Oliveira Souza.
Trabalho de Conclusão de Curso (graduação)-Universidade Federal Fluminense, Instituto de Matemática e Estatística, Niterói, 2023.

1. Modelo de Regressão Poisson. 2. Google Ads. 3. Inferência Bayesiana. 4. Produção intelectual. I. Souza, Mariana Albi de Oliveira, orientadora. II. Universidade Federal Fluminense. Instituto de Matemática e Estatística. III. Título.

CDD - XXX

Bibliotecário responsável: Debora do Nascimento - CRB7/6368

Resumo

Com as medidas restritivas para reduzir a contaminação da COVID-19, muitos empreendedores e/ou empresas de todos os portes foram afetadas e tiveram que buscar novas formas de trabalhar e atrair clientes. Para inovar e atrair clientes, essas empresas buscaram utilizar ferramentas digitais. No dia a dia são realizadas bilhões de pesquisas no Google e há uma popularização na utilização das ferramentas do Google. Essas ferramentas tem mostrado um grande potencial para o crescimentos das ações de marketing. Utilizando a ferramenta Google Ads as empresas podem usar anúncios para atrair potenciais clientes, divulgando e impulsionando o seu negócio. Este trabalho tem o objetivo de entender o comportamento do perfil do consumidor de uma empresa após o início da utilização da ferramenta Google Ads, ou seja, após serem impactadas pelos anúncios produzidos e divulgados pela ferramenta, além de modelar o número de cliques em anúncios, com base em informações dos usuários e nos parâmetros definidos para a veiculação do anúncio. Com esse intuito, foi utilizado um Modelo de Regressão Poisson sob o ponto de vista da Inferência Bayesiana, cujas estimativas dos parâmetros mostram que o sexo não é uma característica determinante no interesse dos usuários impactados pelos anúncios e que o principal público alcançado pelos anúncios em questão é da faixa etária 55 a 64 anos, enquanto o público menos alcançado é da faixa etária de 18 a 24 anos. Estas estimativas fornecem indicativos de como os investimentos da empresa nos anúncios podem ser otimizados; ou seja, a partir destas informações a empresa pode decidir restringir seus anúncios ao perfil de usuários mais interessado em seus anúncios, economizando recursos, ou redirecionar seus recursos para atrair usuários de um perfil não alcançado.

Palavras-chave: Modelo de Regressão Poisson. Google Ads. Inferência Bayesiana.

Sumário

Lista de Figuras

Lista de Tabelas

1 Introdução

1.1 Motivação

A pandemia da doença COVID-19 gerou uma grande crise sanitária mundial e medidas foram criadas para tentar conter o avanço da pandemia. Como a COVID-19 é uma doença com uma taxa de infecção muito alta, as principais medidas para conter esse avanço foram o distanciamento social e a utilização de máscaras.

Com o distanciamento social, muitas empresas foram prejudicadas, pois não tinham mais a circulação de pessoas em suas lojas físicas, gerando um impacto muito grande para os empreendedores. A solução tomada foi tentar inovar e, com a utilização da internet, alguns negócios investiram no uso das ferramentas digitais para atrair e divulgar suas empresas para os clientes. Seguindo o que diz SEBRAE (2021), 86% das empresas decidiram inovar, operando de uma forma diferente e utilizando ferramentas digitais. Dentre essas ferramentas digitais, a que obteve maior exito foi o WhatsApp para negócios, seguido pela ferramenta de propaganda do Google.

O Google Ads é uma ferramenta de marketing digital, que cria anúncios para divulgar empresas, mostrando aos clientes seus diferenciais e alcançando clientes quando eles pesquisam produtos (ou termos relacionados) que a empresa oferece. Na criação de um anúncio, o empreendedor configura a sua campanha definindo palavras chaves que descrevem o produto (utilizadas como termos de busca nas pesquisas dos usuários), escolhendo o público alvo que será impactado pelos anúncios através de configurações como a localidade que deseja atuar com os seus produtos/serviços e/ou características do grupo de clientes que receberá os anúncios ao realizar pesquisas feitas no Google; definindo, por exemplo, sexo ou faixas etárias específicas.

Dentre os recursos oferecidos pela ferramenta, existem objetivos de publicidade que o usuário pode escolher, como: receber mais chamadas na sua empresa, atrair mais visitantes para a sua loja física e/ou direcionar as pessoas para o seu site e decidir se o

anúncio terá alcance global ou local, além de poder definir seu limite de orçamento, só tendo custo se tiver resultados, ou seja, só há custo para o anunciante quando o cliente clica no anúncio para fazer chamadas para sua empresa, acessar seu site ou ver rotas até sua loja.

Os anúncios são exibidos quando as pessoas pesquisam produtos ou serviços semelhantes aos que são oferecidos. Os anúncios podem ser vistos na Pesquisa do Google, no Google Maps e em toda a rede de sites parceiros. A tecnologia inteligente do Google ajuda a encontrar formas para melhorar anúncios e gerar melhores resultados, fornecem relatórios, percepções e dicas para entender seu progresso e otimizar ainda mais os anúncios (GOOGLE, 2023).

Durante a COVID-19, os empreendedores viram no Google Ads uma oportunidade de alcançar novos clientes. Essa ferramenta digital possibilitou que as empresas se adaptassem ao contexto desafiador e procurassem maneiras de continuar atraindo clientes mesmo durante esse período complicado.

Nos relatórios fornecidos pelo Google Ads há dados que permitem a análise descritiva, que por um lado são importantes para entendimento do desempenho do anúncio mas, por outro lado, se faz necessário ter conhecimento prévio por parte do dono do negócio para interpretar tais resultados. O intuito do presente trabalho é fazer mais do que as análises descritivas: é usar estimação de modelos para melhorar as análises dos anúncios e dar mais um indicativo de como configurar a sua campanha.

Outros trabalhos já exploraram os dados gerados pelo Google Ads. Por exemplo, em Martins (2019) foi utilizada a teoria Bayesiana para determinar a forma de encontrar a melhor configuração para a campanha de anúncios no Google Ads. Além desse, Kumar et al. (2015) propõem um modelo para prever as CTR's (taxas de cliques) de anúncios adotando a regressão logística como método.

A classe de modelos lineares generalizados é utilizada em diversas áreas de aplicação. Por exemplo, Habte e Dessu (2023) utilizaram a análise de regressão Poisson para modelar dados de saúde, com objetivo de avaliar a aceitação dos elementos chaves dos serviços de SSR (saúde sexual e reprodutiva) e seus determinantes entre adolescentes residentes nos distritos rurais da zona de Guraghem, no sul da Etiópia. Outro exemplo é o trabalho de Ramos (2000), que utiliza modelos lineares generalizados na previsão de reservas para pagamentos de sinistros. Sob a perspectiva Bayesiana, pode-se citar Toharudin et al. (2020), que utiliza o modelo Poisson para obter a incidência de COVID-19 no oeste de Java, Indonésia.

No presente trabalho, o objetivo é modelar o número de cliques em anúncios de uma empresa que utiliza a ferramenta do Google Ads para impulsionar seu negócio. Quando lidamos com dados de contagem, uma abordagem adequada é utilizar um modelo linear generalizado da família Poisson e uma maneira de estimar o modelo Poisson se dá através da abordagem Bayesiana, que permite incorporar informações a priori sobre os parâmetros do modelo. Mais especificamente, ao utilizar o modelo linear generalizado da família Poisson sob o ponto de vista Bayesiano, espera-se compreender os efeitos das variáveis explicativas, que são: impressões, custo, sexo e faixa etária sobre a quantidade de cliques diária nos anúncios, permitindo uma melhor compreensão sobre o perfil de clientes da empresa e, consequentemente a escolha da melhor configuração da campanha. Além disso, essa abordagem permite fazer previsões sobre a taxa de cliques em novas observações.

1.2 Organização

Esse trabalho está organizado em 4 Capítulos. O primeiro Capítulo trata da motivação do trabalho e uma breve revisão de literatura sobre os temas abordados. O segundo Capítulo contém a apresentação do modelo generalizado no contexto da Inferência Bayesiana, assim como a metodologia aplicada no trabalho. No Capítulo 3, encontram-se os resultados da aplicação dos métodos ao estudo de caso. Por fim, no Capítulo 4, são apresentadas as conclusões do trabalho.

2 Materiais e Métodos

Neste capítulo são discutidos todos os materiais e métodos utilizados no presente trabalho. Inicialmente, são apresentado os Modelos Lineares Múltiplos e sua generalização, os chamados Modelos Lineares Generalizados; seguindo com as seções sobre a estimação do ponto de vista Bayesiano e com os métodos computacionais a serem utilizadas na aplicação apresentada no Capítulo 3.

2.1 Modelos Lineares

É comum na Estatística buscar a explicação de uma variável, chamada variável resposta, por uma ou mais variáveis explicativas, também chamadas de covariáveis. Quando a variável resposta, de fato, pode ser explicada por uma ou mais variáveis explicativas, então a variável resposta tem uma tendência que varia com as covariáveis. Em geral, a notação da variável resposta é dada por Y e da variável explicativa por X ou, no caso de k variáveis explicativas, $X_1, X_2, \ldots, X_k$. Caso a relação entre a resposta e cada uma das variáveis explicativas seja linear, pode-se modelar esta resposta através de um modelo de regressão linear.

O modelo de regressão linear múltipla é descrito da seguinte forma:

$$\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\epsilon} \ , \tag{2.1}$$

onde $\boldsymbol{Y}$ representa o vetor de observações

$$\boldsymbol{Y} = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} ;$$

$\boldsymbol{X}$ é a matriz cujas colunas representam os valores observados para cada uma das k covariáveis

$$\boldsymbol{X} = \begin{bmatrix} X_{1,1} & X_{1,2} & \dots & X_{1,k} \\ X_{2,1} & X_{2,2} & \dots & X_{2,k} \\ \vdots & \vdots & \ddots & \dots \\ X_{n,1} & X_{n,2} & \dots & X_{n,k} \end{bmatrix};$$

$\boldsymbol{\beta}$ é um vetor de parâmetros desconhecidos, chamados de coeficiente de regressão,

$$\boldsymbol{\beta} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{bmatrix}; \ e$$

$\boldsymbol{\epsilon}$ é um vetor de erros, com distribuição normal multivariada com vetor de médias iguais a zero e matriz de covariâncias $\sigma^2 I_n$, onde I_n representa a matriz identidade de dimensão $n \times n$

$$\boldsymbol{\epsilon} = \begin{bmatrix} \epsilon_1 \\ \vdots \\ \epsilon_n \end{bmatrix}.$$

A definição do modelo pode ser resumida em 4 itens, conforme descrito a seguir:

(i) A média de cada Y_i, $i = 1, \dots, n$, é a função linear de $\boldsymbol{X}_i^T = (X_{i,1}, \dots, X_{i,k})$, isto é $E[Y_i] = \beta_1 X_{i,1} + \beta_2 X_{i,2} + \cdots + \beta_k X_{i,k}$. Em particular, se o intercepto for usado, basta supor $\boldsymbol{X}_i^T = (1, \dots, 1)$;

(ii) as variáveis $Y_1, \dots, Y_n$ são independentes;

(iii) a variância de Y_i, $i = 1, \dots, n$, é constante ao longo das observações e denotada por σ^2;

(iv) a distribuição de cada Y_i, $i = 1, \dots, n$, é normal; especificamente, $Y_i \sim N(\boldsymbol{X}_i^T \boldsymbol{\beta}, \sigma^2)$, $i = 1, \dots, n$.

Para entender a interpretação dos parâmetros do modelo, admita o seguinte exemplo. Suponha que uma empresa deseja modelar o salário dos funcionários. As variáveis explicativas poderiam ser o tempo de estudo do funcionário e seu tempo de serviço (ambos em anos). Nesse caso, pode-se assumir um modelo de regressão múltipla descrito como:

$$Y_i = \boldsymbol{X}_i^T \boldsymbol{\beta} + \epsilon_i = \beta_1 X_{i,1} + \beta_2 X_{i,2} + \beta_3 X_{i,3} + \epsilon_i, \ i = 1, \dots, n,$$

onde

- Y_i é o salário do i-ésimo funcionário;
- $X_{i,1} = 1$, para $i = 1, \ldots, n$ (de forma que β_1 representa um efeito comum no salário de todos os funcionários - um intercepto);
- $X_{i,2}$ representa o tempo de estudo (em anos) do i-ésimo funcionário;
- $X_{i,3}$ representa o tempo de serviço (em anos) do i-ésimo funcionário;
- $\epsilon_i, i = 1, \ldots, n$, são erros aleatórios independentes e identicamente distribuídos, com distribuição normal com média zero e variância constante.

Nesse caso, pode-se interpretar os coeficientes β_2 e β_3 da seguinte forma: β_2 é o incremento no salário do funcionário ao aumentar em 1 ano seu tempo de estudo e β_3 é o acréscimo ao salário do funcionário ao aumentar em 1 ano o seu tempo de serviço.

2.1.1 Estimação pontual dos parâmetros

Usualmente, a estimação dos coeficientes de regressão é feita pelo Método de Mínimos Quadrados. É uma técnica de otimização matemática que busca encontrar o melhor ajuste para o conjunto de dados tentando minimizar a soma dos quadrados das diferenças entre os valores observados e suas respectivas médias. Em particular, para um modelo de regressão linear múltiplo, esta soma toma a forma $SQ(\boldsymbol{\beta}) = \sum_{i=1}^{n} \epsilon_i^2 = \sum_{i=1}^{n} (Y_i - \boldsymbol{X}_i^T \boldsymbol{\beta})^2$ e pode-se provar que o valor que minimiza a soma $SQ(\boldsymbol{\beta})$ é

$$\hat{\boldsymbol{\beta}} = (\boldsymbol{X}^T \boldsymbol{X})^{-1} \boldsymbol{X}^T \boldsymbol{Y}.$$

Para mais detalhes veja Jorgensen (2019).

Uma restrição que existe na aplicação do Método dos Mínimos Quadrados nesta classe de modelos é ele não estima a variância, uma vez que esse parâmetro não está relacionado à média das observações. Uma alternativa, nesse caso, é utilizar estimadores de máxima verossimilhança.

A ideia do método da máxima verossimilhança é escolher como estimativa para o vetor de parâmetros desconhecidos o valor que torna a amostra observada a mais provável possível. Para isso, deve-se maximizar a função de verossimilhança do modelo.

No caso do modelo linear múltiplo, cujas observações seguem a distribuição normal, a função verossimilhança, para um dado vetor de observações $\boldsymbol{y} = (y_1, \ldots, y_n)^T$ de $\boldsymbol{Y}$, tem a seguinte forma:

$$l(\boldsymbol{\beta}, \sigma^2; \boldsymbol{y}) = \prod_{i=1}^{n} \left[\frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{ -\frac{1}{2\sigma^2}(y_i - \boldsymbol{X}_i^T\boldsymbol{\beta})^2 \right\} \right] \tag{2.2}$$

Os estimadores para $\boldsymbol{\beta}$ e σ^2 são obtidos maximizando a função (2.2). Nesse caso, o estimador $\hat{\boldsymbol{\beta}}$ de $\boldsymbol{\beta}$ coincide com o que é obtido utilizando o Método dos Mínimos Quadrados, ou seja,

$$\hat{\boldsymbol{\beta}} = (\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{Y}, \tag{2.3}$$

e o estimador $\hat{\sigma}^2$ da variância σ^2 é da seguinte forma:

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(\boldsymbol{Y} - \boldsymbol{X_i}^T\boldsymbol{\beta})^2. \tag{2.4}$$

Novamente, para mais detalhes veja Jorgensen (2019).

Uma restrição importante no modelo de regressão linear múltipla é a hipótese de que a distribuição das observações é normal. No mundo real há muitas variáveis que não seguem essa distribuição. Para retirar essa limitação, Nelder e Wedderburn (1972) propuseram novos caminhos abordando uma outra metodologia chamada modelos lineares generalizados, que é descrita na próxima seção.

2.2 Modelos Lineares Generalizados

Conforme descrito em Dobson e Barnett (2018), o modelo linear generalizado (MLG) é definido em termos de um conjunto de variáveis aleatórias independentes $Y_1, ..., Y_n$. As variáveis resposta podem seguir uma distribuição diferente da normal, podendo, por exemplo, ser categóricas ao invés de contínuas. Além disso, a relação entre a média da variável resposta e as covariáveis não precisa ser linear. A variável resposta segue uma distribuição na família exponencial e as seguintes propriedades são assumidas para essa classe de modelos.

- A distribuição de cada Y_i, $i = 1, ...n$, pertence à família exponencial de distribuições da forma canônica e depende de um único parâmetro θ_i. Nesse caso, a função de probabilidade/função de densidade de probabilidade para cada valor y_i de Y_i

assumirá a forma

$$p(y_i|\theta_i) = \exp[y_i b_i(\theta_i) + c_i(\theta_i) + d_i(y_i)], \tag{2.5}$$

onde $b_i(\cdot)$, $c_i(\cdot)$ e $d_i(\cdot)$ são funções reais;

- Todos os Y_i's seguem a mesma distribuição, sendo possível retirar os subscritos em b, c e d. Dessa forma, a função de probabilidade/função de densidade de probabilidade conjunta de $\boldsymbol{Y}$ no ponto $\boldsymbol{y} = (y_1, ..., y_n)^T$ é

$$\begin{aligned} p(\boldsymbol{y}|\boldsymbol{\theta}) &= \prod_{i=1}^{n} \exp\left[y_i b(\theta_i) + c(\theta_i) + d(y_i)\right] \\ &= \exp\left[\sum_{i=1}^{n} y_i b(\theta_i) + \sum_{i=1}^{n} c(\theta_i) + \sum_{i=1}^{n} d(y_i)\right], \end{aligned} \tag{2.6}$$

onde $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_n)^T$;

- A média de Y_i, $i = 1, \ldots, n$ é denotada por $E(Y_i) = \mu_i$ e é alguma função de θ_i. Para um modelo linear generalizado existe uma transformação de μ_i tal que

$$g(\mu_i) = \boldsymbol{X}_i^T \boldsymbol{\beta},$$

onde g é uma função monótona e diferenciável chamada de função de ligação e $\boldsymbol{X}_i^T$ e $\boldsymbol{\beta}$ estão definidos da mesma forma que na equação (2.1).

Várias distribuições estão incluídas na família exponencial, podendo ser escritas da forma (2.6). Como exemplos, pode-se citar as distribuições normal, normal inversa, gama, Poisson, Bernoulli, entre outras. Isto significa grande flexibilidade para modelar diferentes tipos de fenômenos utilizando MLG. Nas Seções 2.2.2 e 2.2.3 são apresentados dois casos particulares dessa classe de modelos.

2.2.1 Estimação pontual dos parâmetros

Para o modelo linear generalizado (MLG), novamente a estimativa dos coeficientes de regressão podem ser encontradas maximizando a função verossimilhança, ou, equivalentemente, a função de log-verossimilhança. Para um dado vetor de observações $\boldsymbol{y} = (y_1, \ldots, y_n)^T$ de $\boldsymbol{Y}$, a função de log-verossimilhança, denotada por $L(\boldsymbol{\theta}; \boldsymbol{y})$, assume a forma:

$$L(\boldsymbol{\theta}; \boldsymbol{y}) = \sum_{i=1}^{n} y_i b(\theta_i) + \sum_{i=1}^{n} c(\theta_i) + \sum_{i=1}^{n} d(y_i). \tag{2.7}$$

As distribuições da família exponencial satisfazem a certas condições de regularidade que garantem que o máximo global da log-verossimilhança é dado unicamente pela solução das equações $\dfrac{\partial L(\boldsymbol{\theta};\boldsymbol{y})}{\partial \boldsymbol{\theta}} = \mathbf{0}$ (em que cada derivada parcial de $L(\boldsymbol{\theta};\boldsymbol{y})$ é igual a zero). No entanto, para resolver o sistema, em geral é necessário utilizar aproximações numéricas. Para mais detalhes veja Dobson e Barnett (2018).

Nesse trabalho, quando necessário, estas aproximações serão feitas utilizando métodos numéricos que estão implementados no *software* estatístico R (R Core Team, 2022).

2.2.2 Modelo de Regressão Logística

O modelo regressão logística é um caso particular do modelo linear generalizado, no qual a variável reposta é binária. Mais especificamente, se $Y_i \sim Bernoulli(\pi_i)$, com $Y_i = 1$ indicando a observação de um sucesso e $Y_i = 0$ a ocorrência de um fracasso, supondo uma amostra de n observações independentes com $P(Y_i = 1) = \pi_i$, para um dado vetor de observações $\boldsymbol{y} = (y_1, \ldots, y_n)^T$ de $\boldsymbol{Y}$,

$$p(\boldsymbol{y}|\boldsymbol{\pi}) = \prod_{i=1}^{n} \pi_i^{y_i}(1-\pi_i)^{1-y_i} = \exp\left[\sum_{i=1}^{n} y_i \log\left(\frac{\pi_i}{1-\pi_i}\right) + \sum_{i=1}^{n} \log(1-\pi_i)\right], \tag{2.8}$$

o que indica que essa distribuição pertence à família exponencial (basta comparar com a equação (2.6)).

A função de ligação para esse modelo é dada por:

$$\log\left(\frac{\pi_i}{1-\pi_i}\right) = \boldsymbol{X}_i^T\boldsymbol{\beta}. \tag{2.9}$$

O termo $\log\left(\dfrac{\pi_i}{1-\pi_i}\right)$ tem uma interpretação natural como o logarítmo das chances, sendo as chances dada por $\exp(\boldsymbol{X}_i^T\boldsymbol{\beta})$.
Cada β_j, $j = 1, \ldots, k$, mede o efeito do aumento em uma unidade de $X_{i,j}$, mantendo os valores das outras covariáveis constantes. Seguindo o modelo logístico pode-se obter a probabilidade a favor do evento de interesse da i-ésima unidade, representada na expressão abaixo:

$$\pi_i = \frac{\exp(\boldsymbol{X}_i^T\boldsymbol{\beta})}{1+\exp(\boldsymbol{X}_i^T\boldsymbol{\beta})}. \tag{2.10}$$

Portanto, a função de verossimilhança em função de $\boldsymbol{\beta}$ seguindo a equação (2.8) e substituindo o valor de π_i dado na expressão (2.10), para um dado vetor de observações

$\boldsymbol{y} = (y_1, \ldots, y_n)^T$ de $\boldsymbol{Y}$ toma a forma

$$l(\boldsymbol{\beta}; \boldsymbol{y}) = \exp\left[\sum_{j=1}^{n} y_i \log\left(\frac{\pi_i}{1-\pi_i}\right) + \sum_{i=1}^{n} \log(1-\pi_i)\right]$$

$$= \exp\left[\sum_{i=1}^{n} y_i \boldsymbol{X}_i^T \boldsymbol{\beta} - \log(1 + \exp(\boldsymbol{X}_i^T \boldsymbol{\beta}))\right].$$

Os estimadores para $\boldsymbol{\beta}$, e consequentemente para $\pi_i = g^{-1}(\boldsymbol{X}_i^T \boldsymbol{\beta})$, podem ser obtidas maximizando essa função com relação às componentes do vetor $\boldsymbol{\beta}$, ou equivalentemente, maximizando a função de log-verossimilhança

$$L(\boldsymbol{\beta}; \boldsymbol{y}) = \sum_{i=1}^{n} y_i \boldsymbol{X}_i^T \boldsymbol{\beta} - \log(1 + \exp(\boldsymbol{X}_i^T \boldsymbol{\beta})).$$

2.2.3 Modelo de Regressão Poisson

O modelo de regressão Poisson é outro caso do modelo linear generalizado, no qual a variável resposta corresponde a um dado de contagem. São exemplos de variáveis resposta que podem ser modeladas pela distribuição Poisson dados de contagem como: número de acidentes no trânsito, quantidade de casos de uma doença em certa região e o número de cliques em anúncios de uma campanha de marketing, que é o caso tratado nesse trabalho.

Sejam $Y_1, \ldots, Y_n$ variáveis independentes tais que $Y_i \sim Poisson(\mu_i)$, $i = 1, \ldots, n$. Nesse caso o parâmetro μ_i representa a frequência média ou a taxa de ocorrência desse evento e tanto o valor esperado de Y_i como sua variância são representados por μ_i; ou seja,

$$E(Y_i) = Var(Y_i) = \mu_i. \tag{2.11}$$

Para uma dada amostra observada $\boldsymbol{y} = (y_1, ..., y_n)^T$ de $\boldsymbol{Y} = (Y_1, ..., Y_n)^T$,

$$p(\boldsymbol{y}|\boldsymbol{\mu}) = \prod_{i=1}^{n} \frac{\mu_i^{y_i} \exp(-\mu_i)}{y_i!} = \exp\left[\sum_{i=1}^{n} (\log(\mu_i) y_i - \log(y_i!) - \mu_i)\right], \tag{2.12}$$

onde $\boldsymbol{\mu} = (\mu_1, \mu_2, \ldots, \mu_n)^T$, observa-se, comparando com a equação (2.6), que a distribuição Poisson pertence a família exponencial.

A função de ligação natural nesse caso é a logarítmica

$$log(\mu_i) = \boldsymbol{X}_i^T \boldsymbol{\beta}, \tag{2.13}$$

indicando que a dependência de μ_i em relação às covariáveis é expressa como

$$\mu_i = \exp(\boldsymbol{X}_i^T \boldsymbol{\beta}),$$

com $\boldsymbol{\beta}$ e $\boldsymbol{X}_i^T$ definidos como na Seção 2.1.

Substituindo (2.13) na equação (2.12), obtemos a seguinte função de verossimilhança para os parâmetros do modelo

$$l(\boldsymbol{\beta}; \boldsymbol{y}) = \exp\left[\sum_{i=1}^{n} (\boldsymbol{X}_i^T \boldsymbol{\beta} y_i - \log(y_i!) - \exp(\boldsymbol{X}_i^T \boldsymbol{\beta})\right] \tag{2.14}$$

para um dado valor observado $\boldsymbol{y}$ de $\boldsymbol{Y}$.

Note que, para uma dada configuração de $\boldsymbol{X}_i^T$, digamos $\boldsymbol{x}_i^T = (x_{i,1}, x_{i,2}, \ldots, x_{i,k})$,

$$E(Y_i|\boldsymbol{x}_i^T) = \exp(\boldsymbol{x}_i^T \boldsymbol{\beta}).$$

Por outro lado, se ocorre o acréscimo de 1 unidade na j-ésima componente de $\boldsymbol{x}_i^T$ (mantendo todas as demais constantes), isto é, para $\boldsymbol{w}_i^T = (x_{i,1}, x_{i,2}, \ldots, x_{i,j}+1, \ldots, x_{i,k})$, temos

$$\frac{E(Y_i|\boldsymbol{w}_i^T)}{E(Y_i|\boldsymbol{x}_i^T)} = \frac{\exp(\boldsymbol{w}_i^T \boldsymbol{\beta})}{\exp(\boldsymbol{x}_i^T \boldsymbol{\beta})} = \frac{\exp(\beta_1 x_{i,1} + \beta_2 x_{i,2} + \cdots + \beta_j (x_{i,j}+1) + \cdots + \beta_k x_{i,k})}{\exp(\beta_1 x_{i,1} + \beta_2 x_{i,2} + \cdots + \beta_j x_{i,j} + \cdots + \beta_k x_{i,k})} = \exp(\beta_j). \tag{2.15}$$

A quantidade definida na equação (2.15) é chamada de razão de taxas e pode ser interpretada da seguinte forma: quando o valor da variável explicativa X_j aumenta em uma unidade, a média das observações aumenta (ou diminui) $\exp(\beta_j)$ vezes, ou seja, resulta em um efeito multiplicativo de $\exp(\beta_j)$ na taxa μ.

Em particular, quando a variável explicativa é uma variável categórica binária, representada por uma variável *dummy*, a interpretação para razão de taxas segue-se da seguinte maneira: supondo que $X_j = 1$ indica a presença de um dado fator e $X_j = 0$ indica a ausência desse fator, a razão de taxas $\exp(\beta_j)$ indica o efeito da presença do fator na média das observações, em comparação à sua ausência. Mais especificamente, a taxa de ocorrência na presença do fator X_j é $\exp(\beta_j)$ vezes a taxa de ocorrência na ausência do fator.

Essa interpretação pode ser naturalmente estendida para o caso de variáveis categóricas com mais categorias, através da criação de variáveis auxiliares *dummies*. Para utilizar as variáveis *dummies*, é comum separar uma categoria de referência da variável

categórica e criar *dummies* adicionais para representar as demais categorias (assumindo valor 1 se a observação pertencer à categoria correspondente). A categoria de referência é utilizada como base de comparação e seu efeito é capturado pelo intercepto (considerando o valor 0 (zero) em todas as *dummies* associadas à variável categórica em questão). Com isso pode-se interpretar o efeito de cada categoria em relação à categoria de referência através da exponencial do coeficiente associado à sua respectiva *dummy*.

2.3 Teste de Wald

O teste de Wald permite avaliar quais variáveis explicativas do modelo tem efeito estatisticamente significativo, ao nível de significância α. Conforme descrito no Capítulo 5 de Dobson e Barnett (2018), pode-se afirmar que o estimador de máxima verossimilhança $\hat{\beta}_j$ de β_j, $j = 1, \ldots, k$, é tal que $\hat{\beta}_j \sim N(\beta_j, Var(\hat{\beta}_j))$. Através disso tem-se a seguinte estatística de teste:

$$Z_j = \frac{\hat{\beta}_j - \beta_j}{\sqrt{Var(\hat{\beta}_j)}} \sim N(0,1). \tag{2.16}$$

Para cada $j = 1, \ldots, k$, as hipótese do teste a serem testadas são definidas, como

$$\begin{cases} H_0 : \beta_j = 0 \\ H_1 : \beta_j \neq 0, \end{cases}$$

em que H_0 indica que não existe influência significativa entre a variável resposta e a variável explicativa X_j e H_1 supõe a existência de associação significativa entre a variável resposta e a variável explicativa X_j.

Sob H_0, a estatística de teste observada Z_j será definida como:

$$Z_j = \frac{\hat{\beta}_j}{\sqrt{Var(\hat{\beta}_j)}}. \tag{2.17}$$

A tomada de decisão mais utilizada é a partir do p-valor do teste de Wald, calculado da forma:

$$p\text{-valor} = P(Z \geq |Z_j|).$$

Rejeita-se H_0 se o p-valor é menor ou igual ao nível de significância α; caso contrário, não há evidências para rejeitar H_0 ao nível de significância de α. Para mais detalhes, consulte Dobson e Barnett (2018).

2.4 Inferência Bayesiana

Quando o objetivo é fazer inferência estatística, em geral usa-se apenas as informações que vêm das observações para inferir e chegar a uma solução. Mas se for adicionado um conhecimento subjetivo prévio do pesquisador, através da especificação de uma distribuição a priori para os parâmetros do modelo, é possível combinar as duas fontes de informações utilizando o Teorema de Bayes (MIGON; GAMERMAN; LOUZADA, 2014).

Teorema 2.1 (**Teorema de Bayes**). *Seja* $\boldsymbol{y} = (y_1, \ldots, y_n)^T$ *um vetor de observações uma variável aleatória* $\boldsymbol{Y}$ *cuja distribuição é descrita por um vetor de parâmetro* $\boldsymbol{\theta}$*. Nesse caso, a distribuição a posteriori de* $\boldsymbol{\theta}$ *dado* $\boldsymbol{y}$ *pode ser obtida por*

$$p(\boldsymbol{\theta}|\boldsymbol{y}) = \frac{p(\boldsymbol{y}|\boldsymbol{\theta})p(\boldsymbol{\theta})}{p(\boldsymbol{y})} \propto l(\boldsymbol{\theta}; \boldsymbol{y})p(\boldsymbol{\theta}) \tag{2.18}$$

onde

- $p(\boldsymbol{\theta}|\boldsymbol{y})$ *é a distribuição a posteriori do vetor de parâmetros* $\boldsymbol{\theta}$*;*
- $p(\boldsymbol{y})$ *é a distribuição marginal de* $\boldsymbol{y}$*;*
- $l(\boldsymbol{\theta}; \boldsymbol{y})$ *é a função de verossimilhança de* $\boldsymbol{\theta}$ *para cada vetor de observações* $\boldsymbol{y}$ *de* $\boldsymbol{Y}$ *(tendo a mesma expressão do modelo observacional* $p(\boldsymbol{y}|\boldsymbol{\theta})$*);*
- $p(\boldsymbol{\theta})$ *é a distribuição a priori do vetor de parâmetros* $\boldsymbol{\theta}$*.*

Em muitos casos, descobrir qual a forma analítica da distribuição a posteriori obtida pela expressão (2.18) é uma tarefa muito complexa e, para conseguir prosseguir com as análises, são utilizados algoritmos que permitem entender melhor o comportamento das distribuições a posteriori. Nesse trabalho, os algoritmos abordados com esse objetivo fazem parte dos chamados Métodos de Monte Carlo via Cadeias de Markov e servem para gerar amostras de distribuições de interesse, facilitando a estimação dos parâmetros (MIGON; GAMERMAN; LOUZADA, 2014).

2.4.1 Métodos de Monte Carlo via Cadeias de Markov

Monte Carlo via Cadeias de Markov (MCMC) é uma ampla classe de ferramentas computacionais para geração de amostras e aproximação de integrais. Na análise Bayesiana, os algoritmos MCMC são usados principalmente para simular amostras que aproximam a distribuição a posteriori do vetor de parâmetros de interesse. Exemplos de algoritmos

MCMC incluem o Metropolis-Hastings, o Amostrador de Gibbs e o Monte Carlo Hamiltoniano.

A ideia central nos algoritmos MCMC é gerar uma sequência de valores de forma iterativa, permitindo que a cada passo o algoritmo aprenda com o passo anterior (conforme ocorre nas Cadeias de Markov) (GAMERMAN; LOPES, 2006).

2.4.2 Algoritmo Metropolis-Hastings

Metropolis et al. (1953), desenvolveram um algoritmo para gerar amostras baseados no método de aceitação e rejeição. Conforme descrito em Ehlers (2003), suponha que deseja-se gerar um valor $\boldsymbol{\theta}$ de acordo com um modelo descrito por $f(\boldsymbol{\theta})$, porém não se sabe gerar diretamente desta distribuição. Nesse caso, um valor é gerado de uma distribuição auxiliar $q(\boldsymbol{\theta})$ e aceita-se o valor gerado com uma certa probabilidade que depende da distribuição de interesse $f(\boldsymbol{\theta})$. Repetindo esse processo um grande número de vezes, esse mecanismo de correção garante a convergência da cadeia de Markov de valores amostrados para a distribuição de equilíbrio de interesse.

Suponha que numa dada iteração do algoritmo, a cadeia de Markov esteja no estado $\boldsymbol{\theta}^c$ e um valor $\boldsymbol{\theta}^p$ é gerado de uma distribuição proposta $q(\cdot|\boldsymbol{\theta}^c)$ (que pode depender do estado atual da cadeia $\boldsymbol{\theta}^c$). O novo valor $\boldsymbol{\theta}^p$ é aceito de acordo com a probabilidade abaixo:

$$\alpha(\boldsymbol{\theta}^c, \boldsymbol{\theta}^p) = min\left\{1, \frac{f(\boldsymbol{\theta}^p)q(\boldsymbol{\theta}^c|\boldsymbol{\theta}^p)}{f(\boldsymbol{\theta}^c)q(\boldsymbol{\theta}^p|\boldsymbol{\theta}^c)}\right\}, \tag{2.19}$$

onde f é a distribuição de interesse. Note que só é preciso conhecer f parcialmente (a menos de constantes), pois nesse caso a probabilidade não se altera.

Seguem os passos do algoritmo:

1. Inicialize o contador de iterações $t = 0$ e especifique um valor inicial $\boldsymbol{\theta}^{(t)}$;
2. gere um novo valor $\boldsymbol{\theta}^p$ da distribuição $q(.|\boldsymbol{\theta}^{(t)})$;
3. calcule a probabilidade de aceitação $\alpha(\boldsymbol{\theta}^{(t)}, \boldsymbol{\theta}^p)$ e gere $u \sim U(0, 1)$;
4. se $u < \alpha(\boldsymbol{\theta}^{(t)}, \boldsymbol{\theta^p})$, então aceite o novo valor e faça $\boldsymbol{\theta}^{(t+1)} = \boldsymbol{\theta}^p$, caso contrário rejeite e faça $\boldsymbol{\theta}^{(t+1)} = \boldsymbol{\theta}^{(t)}$;
5. incremente o contador de t para $t + 1$ e volte ao passo 2 até obter a convergência.

Sob o ponto de vista da inferência Bayesiana, é comum que o interesse esteja na distribuição a posteriori do parâmetro. Nesse caso, o algoritmo é utilizado para gerar amostras da distribuição a posteriori com forma desconhecida tomando $f(\boldsymbol{\theta}) = p(\boldsymbol{\theta}|\boldsymbol{y})$.

2.4.3 Amostrador de Gibbs

A ideia do amostrador de Gibbs desenvolvido por Gelman, Gilks e Roberts (1997) é dividir o vetor $\boldsymbol{\theta}$, com distribuição conjunta $f(\boldsymbol{\theta})$, de forma que cada parte $\boldsymbol{\theta_i}$ de $\boldsymbol{\theta}$ possa ser gerada de sua distribuição condicional completa. A cadeia nesse algoritmo sempre irá se mover para um novo valor pois não existe mecanismo de aceitação-rejeição. As transições de um estado para outro são feitas de acordo com as distribuições condicionais completas $f(\boldsymbol{\theta}_i|\boldsymbol{\theta}_{-i})$, onde $\boldsymbol{\theta}_{-i} = (\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_{i-1}, \boldsymbol{\theta}_{i+1}, \ldots, \boldsymbol{\theta}_d)^T$. Cada uma das componentes $\boldsymbol{\theta}_i$ pode ser uni ou multidimensional. A distribuição condicional completa é a distribuição da i-ésima componente de $\boldsymbol{\theta}$ condicionada em todas as outras componentes, sendo obtida a partir da distribuição conjunta pela seguinte equação

$$f(\boldsymbol{\theta}_i|\boldsymbol{\theta}_{-i}) = \frac{f(\boldsymbol{\theta})}{\int f(\boldsymbol{\theta}) d\boldsymbol{\theta}_i};$$

ou seja, para obter o núcleo da distribuição condicional completa de $\boldsymbol{\theta_i}$ basta pegar os termos da distribuição conjunta de $\boldsymbol{\theta}$ que dependem de $\boldsymbol{\theta}_i$.

Se as distribuições condicionais completas forem completamente conhecidas, então o amostrador de Gibbs é definido pelo seguinte algoritmo:

1. Inicialize o contador de iterações da cadeia $t = 0$ e especifique valores iniciais $\boldsymbol{\theta}^{(t)} = (\boldsymbol{\theta}_1^{(t)}, \ldots, \boldsymbol{\theta}_d^{(t)})^T$;

2. obtenha um novo valor de $\boldsymbol{\theta}^{(t+1)}$ a partir de $\boldsymbol{\theta}^{(t)}$ através da geração sucessiva dos valores

$$\begin{aligned}
\boldsymbol{\theta}_1^{(t+1)} &\sim f(\boldsymbol{\theta}_1|\boldsymbol{\theta}_2^{(t)}, \boldsymbol{\theta}_3^{(t)}, \ldots, \boldsymbol{\theta}_d^{(t)}) \\
\boldsymbol{\theta}_2^{(t+1)} &\sim f(\boldsymbol{\theta}_2|\boldsymbol{\theta}_1^{(t+1)}, \boldsymbol{\theta}_3^{(t)}, \ldots, \boldsymbol{\theta}_d^{(t)}) \\
&\vdots \\
\boldsymbol{\theta}_d^{(t+1)} &\sim f(\boldsymbol{\theta}_d|\boldsymbol{\theta}_1^{(t+1)}, \boldsymbol{\theta}_2^{(t+1)}, \ldots, \boldsymbol{\theta}_{d-1}^{(t+1)});
\end{aligned}$$

3. incremente o contador de t para $t + 1$ e retorne ao passo 2 até obter convergência.

Caso alguma(s) condicional(is) completa(s) não seja completamente conhecida(s), não sabendo-se gerar diretamente dela(s), as gerações das componentes do passo 2 do algo-

ritmo podem ser feitas com o auxílio do algoritmo Metropolis-Hastings. Esta variação do algoritmo é conhecida como amostrador de Gibbs com passos de Metropolis-Hastings.

Novamente, tendo interesse em gerar uma amostra da distribuição a posteriori do vetor de parâmetros $\boldsymbol{\theta}$, basta tomar $f(\boldsymbol{\theta}) = p(\boldsymbol{\theta}|\boldsymbol{y})$. Para mais detalhes, consulte Ehlers (2003).

2.5 Monte Carlo Hamiltoniano

No algoritmo Metropolis-Hastings, geralmente são necessárias muitas iterações para explorar áreas de maior densidade a posteriori. O algoritmo Monte Carlo Hamiltoniano (HMC) melhora a eficiência do algoritmo de Metropolis-Hasting empregando um esquema de geração de proposta guiada que usa o gradiente do log da distribuição a posteriori para direcionar a cadeia de Markov para regiões de maior densidade a posteriori, onde a maioria das amostras é gerada. O HMC aparece como uma alternativa na análise Bayesiana devido às limitações existentes no algoritmo de Metropolis-Hastings e no método de Gibbs, ambos com baixas taxas de aceitação (GELMAN; GILKS; ROBERTS, 1997). De acordo com Pérez (2022), uma cadeia HMC bem ajustada aceitará propostas a uma taxa muito maior do que o algoritmo Metropolis-Hastings tradicional.

Nesse trabalho, estimamos os parâmetros do modelo de interesse utilizando o ponto de vista Bayesiano usando esse tipo de algoritmo. Para isso, fazemos uso do *software* R, com o auxílio da função **stan_glm** do pacote **rstanarm** (GOODRICH et al., 2022).

2.6 Teste de Convergência

Para diagnosticar convergência, pode-se olhar para o traço das cadeias de Markov formadas pelos valores amostrados do vetor de parâmetros, obtendo um critério visual de convergência. Quando o gráfico do traço da cadeia apresenta comportamento repetitivo, "visitando" todo o espaço paramétrico, após descartadas as iterações iniciais (conhecidas como *burn-in* ou aquecimento da cadeia), supomos que houve convergência, e os valores amostrados podem ser utilizados para aproximar características da distribuição de interesse. Existem porém testes mais formais para diagnosticar convergência, como, por exemplo, o método de Gelman-Rubin.

Como segue em Gelman e Rubin (1992), o teste de convergência por eles propostos utiliza múltiplas repetições de cadeias para decidir se a convergência foi alcançada

durante a segunda metade de cada amostra. O teste pressupõe que m sequências tenham sido simuladas em paralelo, partindo de diferentes pontos iniciais. Após descartar a primeira metade das iterações como aquecimento (*burn-in*), as m sequências rendem m possíveis inferências sobre os parâmetros. Se essas inferências são bastantes similares, é um indicativo de que a convergência foi alcançada ou está próxima. Com o método de Gelman-Rubin, obtemos uma medida para avaliar a convergência conhecido como *R-hat*, que representa um fator de redução de escala potencial e pode ser interpretado como um fator de diagnóstico da convergência.

O método de Gelman-Rubin compara a variabilidade intra-cadeias com a variabilidade inter-cadeias para calcular a estatística *R-hat*, que é uma estimativa da razão entre a variância média intra-cadeias e a variância média inter-cadeias.

Quando o *R-hat* se aproxima de 1, pode-se concluir que cada um dos m conjuntos de valores gerados ao longo das iterações se aproximam da distribuição de interesse. Nesse trabalho específico, o valor do *R-hat* é obtido na saída da função **stan_glm** do pacote **rstanarm** utilizando o *summary* no modelo. Para um aprofundamento sobre o tema, consulte Gelman e Rubin (1992) e NOGUEIRA (2004).

3 Análise dos Resultados

Neste Capítulo encontram-se os resultados da aplicação dos métodos discutidos no Capítulo 2 ao estudo de caso de interesse. Todas as análises foram realizadas utilizando o software estatístico R (R Core Team, 2022) e o nível de significância definido foi de $\alpha = 5\%$ para todos os testes de hipóteses e nível de confiança/credibilidade de 95% para todos os intervalos apresentados.

3.1 Descrição dos Dados

No presente trabalho, os dados são um estudo de caso obtido através de uma amostra de conveniência de uma empresa que utilizou a ferramenta Google Ads durante um período de agosto de 2021 até maio de 2023. Destacamos aqui que o nome da empresa não será divulgada por questão de segurança dos dados.

Cada linha do banco de dados contém informações sobre o número de cliques diários em anúncios da empresa para cada perfil de usuário configurado para ser impactado pelo anúncio (para cada sexo e faixa etária considerada). Na Tabela 1 é possível ver algumas linhas do banco de dados completo. No primeiro momento o banco de dados continha 183.780 observações de 15 variáveis, porém, após uma análise mais detalhada, pode-se perceber que muitas variáveis eram transformações de outras, sendo retiradas do banco de dados original. Como o interesse estava em verificar o impacto dos anúncios nos usuários do Google, também foram excluídas as linhas em que o número de impressões (que indica o número de vezes em que o anúncio aparece na busca do usuário) do anúncio eram nulas. Estas linhas configuravam momentos em que o anúncio foi pausado pela empresa, não gerando dados sobre a performance dos mesmos. Por último, foram removidas linhas em que as variáveis qualitativas assumiam valor "Desconhecidos" da base de dados, por indicarem dados faltantes. Após esses ajustes, a base de dados foi para 5.429 observações de 5 variáveis. Dessas 5 variáveis, 4 foram consideradas variáveis explicativas, sendo 2 relacionadas ao grupo de usuários que foi impactado pelo anúncio: faixa etária e sexo,

além do número de impressões do anúncio ao qual o grupo de usuários foi submetido no dia e o custo total gerado pelos cliques desses usuários para o anunciante (em reais). A variável resposta considerada é a quantidade de cliques diária nos anúncios expostos na busca do Google para cada perfil de usuário impactado pela campanha.

Tabela 1: Exemplo das 5 primeira linhas do banco de dados utilizados.

Data	Cliques	Impressões	Custo (R$)	Sexo	Faixa Etária
21/08/2021	0	6	0	Feminino	+65
21/08/2021	4	23	3,03	Feminino	55 a 64
21/08/2021	2	14	2,55	Masculino	55 a 64
21/08/2021	21	156	5,28	Masculino	25 a 34
21/08/2021	7	37	1,72	Feminino	18 a 24

3.2 Análise Descritiva

Inicialmente foi realizada uma análise descritiva da variável resposta Y, que corresponde ao número de cliques diários nos anúncios por perfil de usuário, como visto na Figura 1.

A Figura 1 mostra o boxplot do número de cliques diários considerando todos os perfis de usuários impactados pelos anúncios. Pode-se perceber que a mediana do número de cliques está próxima de 3 e que há uma assimetria nos dados. Também pode-se observar que existem valores discrepantes (*outliers*), aumentando a variabilidade da variável resposta.

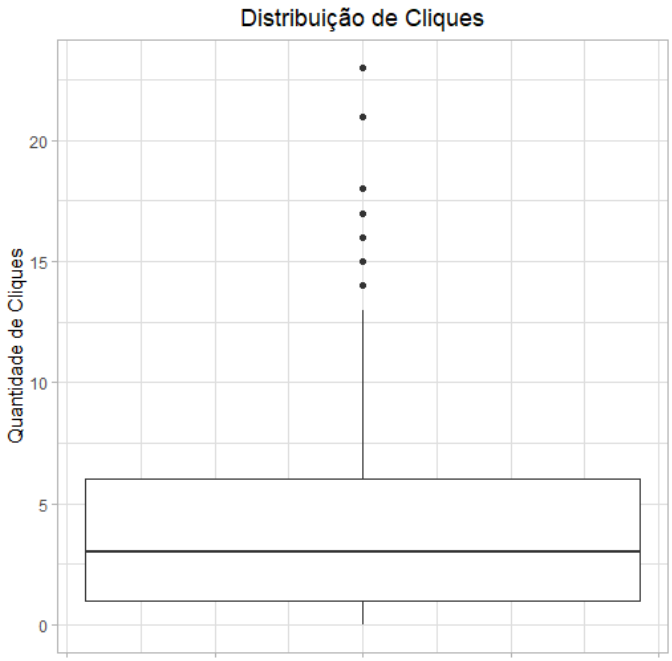

Figura 1: Quantidade de cliques diários considerando todos os perfis de usuários impactados pelos anúncios.

Após esta análise, investigou-se o comportamento das variáveis explicativas, descritas na Tabela 2. Seguindo as análises descritivas, na Tabela 3 observa-se que o mínimo de impressões é igual a 1 (zero impressões quer dizer que o anúncio não estava ativo no momento da busca) e a média é de $65,1$, mas nota-se que as impressões tem um desvio padrão relativamente alta, indicando que os dados tendem a se afastar da média. Para a variável Custo, o menor valor é R$0, indicando que o anúncio não recebeu cliques, com média de R$1,5 e o desvio padrão de R$1,14

Tabela 2: Variáveis explicativas utilizadas na modelagem.

Variável explicativa	Descrição
Impressões	Quantidade de vezes que o anúncio aparece para os clientes de determinado perfil
Custo	Custo total gerado para o criador do anúncio quando há cliques no anúncio (associado a cada perfil de usuário)
Sexo	Sexo dos clientes impactados pelos anúncios
Faixa Etária	Faixa etária dos clientes impactados pelos anúncios

Tabela 3: Medidas resumo para as variáveis explicativas quantitativas Impressões e Custo.

Variável Explicativa	Mín.	1° Quartil	Mediana	Média	3° Quartil	Máx.	Desvio Padrão.
Impressões	1,0	23,0	50,0	65,1	96,0	328,0	52,3
Custo (R$)	0,0	0,4	1,2	1,5	2,3	9,9	1,4

Na Figura 2, observa-se que as variáveis explicativas Custo e Impressões estão relacionadas positivamente com a variável resposta Quantidade de Cliques: quanto maior o custo ou a quantidade de impressões mais cliques são gerados nos anúncios e há uma correlação moderada de 0,64 entre as variáveis Custo e Impressões.

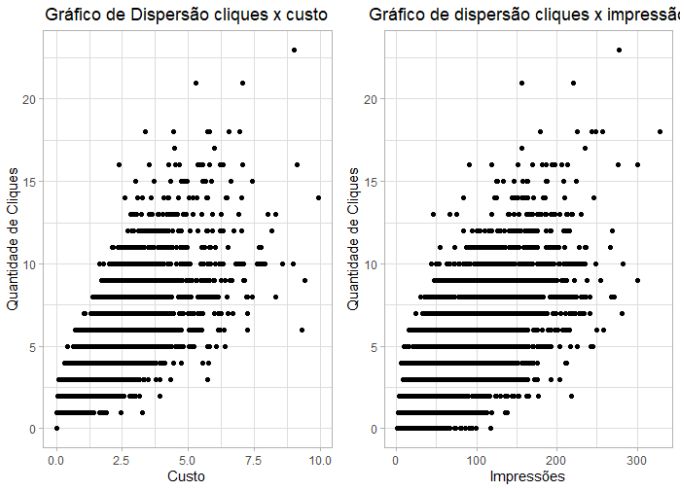

Figura 2: Gráfico de dispersão das variáveis Impressões e Custo *versus* Quantidade de Cliques.

Na Figura 3 temos os boxplots da Quantidade de Cliques por Faixa Etária e por Sexo. Observa-se pela figura que há uma variabilidade maior na quantidade de cliques nas faixas etárias de 35 a 44 anos e de 45 a 54 anos. As medidas de posição maiores nestas faixas etárias também dão indícios de que há um interesse maior nos produtos fornecidos para essas faixas etárias e que os homens clicam mais vezes nos anúncios do que as mulheres.

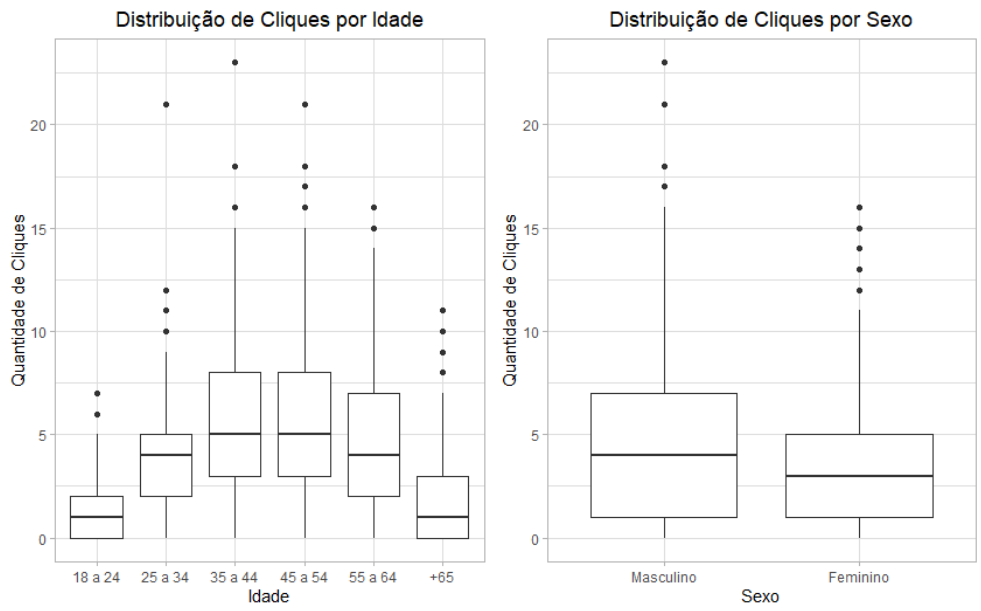

Figura 3: Boxplots da Quantidade de cliques por Faixa Etária e por Sexo.

De fato, após a análise descritiva, parece haver relação entre as variáveis explicativas e a resposta. Para entender melhor estas relações, nas próximas seções, serão ajustados modelos lineares generalizados na família Poisson para medir esses efeitos.

3.3 Estimação dos parâmetros via Modelos Lineares Generalizados com uma covariável

Nesta seção, será analisada a influência marginal das variáveis explicativas apresentadas na Tabela 2 na variável resposta. Com esse objetivo, neste primeiro momento, serão apresentadas as estimativas do modelo linear generalizado da família Poisson para cada covariável separadamente, isto é, tomando uma covariável por vez, com o objetivo de verificar a significância dos efeitos na variável resposta.

As estimativas de todos os parâmetros dos modelos ajustados foram obtidas sob os enfoques clássico e Bayesiano para fins de comparação dos resultados.

3.3.1 Estimação clássica

Para obter as estimativas dos parâmetros dos modelos individuais (considerando apenas uma covariável por vez) sob o enfoque clássico, utilizou-se a função **glm** do *software* R. Em cada um dos modelos considerados utilizou-se um intercepto e, para os modelos com

variáveis explicativas categóricas, utilizou-se variáveis auxiliares *dummies*, retirando-se as categorias "Feminino" e "Mais de 65 anos" como referências para as covariáveis Sexo e Faixa Etária, respectivamente. Os resultados obtidos podem ser vistos nas Tabelas 4, 5, 6 e 7.

É possível notar através das tabelas que em todos os casos o $p-$valor do Teste Wald é menor do que o nível de significância estabelecido de 5%, o que indica que todas as variáveis explicativas têm associação significativa com a variável resposta Quantidade de Cliques. Além disso, apenas a faixa etária de 18 a 24 anos apresenta um efeito negativo, indicando uma redução na taxa de cliques de aproximadamente 50% em usuários que pertencem a esta faixa etária quando comparados à usuários na faixa etária acima dos 65 anos. Na Tabela 7, também é possível notar que, apesar de significativo, o efeito da variável explicativa Impressões é praticamente nulo.

Tabela 4: Estimativas clássicas e p-valor do Teste de Wald para os parâmetros do Modelo Linear Generalizado Poisson considerando apenas a covariável Sexo.

Covariável	Estimativa	Desvio Padrão	p-valor
Intercepto	1,17403	0,01067	$< 0{,}001$
Sexo			
Masculino	0,30995	0,01405	$< 0{,}001$

Tabela 5: Estimativas clássicas e p-valor do Teste de Wald para os parâmetros do Modelo Linear Generalizado Poisson considerando apenas a covariável Faixa Etária.

Covariável	Estimativa	Desvio Padrão	p-valor
Intercepto	0,63442	0,02423	$< 0{,}001$
Faixa Etária			
18 a 24	-0,50920	0,03970	$< 0{,}001$
25 a 34	0,69107	0,02967	$< 0{,}001$
35 a 44	1,11678	0,02789	$< 0{,}001$
45 a 54	1,10593	0,02794	$< 0{,}001$
55 a 64	0,89774	0,02872	$< 0{,}001$

Tabela 6: Estimativas clássicas e p-valor do Teste de Wald para os parâmetros do Modelo Linear Generalizado Poisson considerando apenas a covariável Custo.

Covariável	Estimativa	Desvio Padrão	p-valor
Intercepto	0,64734	0,01119	$< 0{,}001$
Custo	0,34251	0,00337	$< 0{,}001$

Tabela 7: Estimativas clássicas e p-valor do Teste de Wald para os parâmetros do Modelo Linear Generalizado Poisson considerando apenas a covariável Impressões.

Covariável	Estimativa	Desvio Padrão	p-valor
Intercepto	0,56765	0,01250	$< 0{,}001$
Impressões	0,00956	0,00010	$< 0{,}001$

3.3.2 Estimação Bayesiana

Para obter as estimativas dos parâmetros dos modelos individuais sob o ponto de vista Bayesiano, utilizou-se a função **stan_glm** do pacote **rstanarm** do *software* R. Novamente, em cada um dos modelos considerados, utilizou-se um intercepto e, para os modelos com variáveis explicativas categóricas, utilizou-se variáveis auxiliares *dummies*, retirando-se as categorias "Feminino" e "Mais de 65 anos" como referências para as covariáveis Sexo e Faixa Etária, respectivamente.

As distribuições a priori assumidas para todos os parâmetros foram as pré-definidas pela função **stan_glm**, representando distribuições a priori não informativas para os coeficientes dos modelos. Além disso, foram obtidas 4 cadeias para cada um dos parâmetros, tendo cada uma 4.000 iterações, sendo descartadas as 2.000 primeiras como amostra de aquecimento (*burn-in*).

Sob o ponto de vista Bayesiano, antes da obtenção das estimativas dos parâmetros, é importante verificar se há convergência para a distribuição de interesse. Uma forma mais visual de avaliar essa convergência envolve acompanhar as interações do HMC. É importante observar se as trajetórias estão se estabilizando ou verificar a convergência através da estatística *R-hat*, como descrito na Seção 2.6.

Os resultados obtidos sob essa abordagem podem ser vistos nas Tabelas 8, 9, 10 e 11.

Cada uma das tabelas apresenta a estatística *R-hat* de Gelman-Rubin associada as cadeias de cada parâmetros e as estimativas pontuais (médias a posteriori) dos parâmetros. Pode-se notar que a estatística *R-hat* é igual a 1 em todos os casos, indicando convergência. Nata-se também que as estimativas obtidas sob esta abordagem são bem próximas aos resultados da estimação com a abordagem clássica vistos nas Tabelas 4, 5, 6 e 7. Nesse contexto, ambas as abordagens conduzem às mesmas conclusões.

Tabela 8: Estimativas Bayesianas (médias a posteriori) e estatística *R-hat* para os parâmetros do Modelo Linear Generalizado Poisson considerando apenas a covariável Sexo.

Covariável	Estimativa	R-hat
Intercepto	1,17343	1
Sexo		
Masculino	0,31017	1

Tabela 9: Estimativas Bayesianas (médias a posteriori) e estatística *R-hat* para os parâmetros do Modelo Linear Generalizado Poisson considerando apenas a covariável Faixa Etária.

Covariável	Estimativa	R-hat
Intercepto	0,63412	1
Faixa Etária		
18 a 24	-0,50917	1
25 a 34	0,69114	1
35 a 44	1,11672	1
45 a 54	1,10554	1
55 a 64	0,89879	1

Tabela 10: Estimativas Bayesianas (médias a posteriori) e estatística *R-hat* para os parâmetros do Modelo Linear Generalizado Poisson considerando apenas a covariável Custo.

Covariável	Estimativa	R-hat
Intercepto	0,64676	1
Custo	0,34248	1

Tabela 11: Estimativas Bayesianas (médias a posteriori) e estatística *R-hat* para os parâmetros do Modelo Linear Generalizado Poisson considerando apenas a covariável Impressões.

Covariável	Estimativa	R-hat
Intercepto	0,56733	1
Impressão	0,00957	1

3.4 Estimação dos parâmetros via Modelo Linear Generalizado Múltiplo

Tendo em vista que todas as covariáveis tiveram efeito significativo na variável resposta, nesta seção, serão apresentados os resultados obtidos através de um modelo linear generalizado múltiplo, com o objetivo de avaliar o comportamento conjunto das variáveis explicativas e medir seus efeitos na variável resposta. Mais especificamente, vamos utilizar o modelo linear generalizado da família Poisson, conforme descrito na Seção 2.2.3, supondo que Y_i representa o número diário de cliques feitos em anúncios exibidos para usuários de um dado perfil descrito pelas covariáveis apresentadas na Tabela 2.

Conforme foi feito na seção anterior, aqui também utilizou-se um modelo com intercepto e, para as variáveis explicativas categóricas, utilizou-se variáveis auxiliares *dummies*, retirando-se as categorias "Feminino" e "Mais de 65 anos" como referências para as covariáveis Sexo e Faixa Etária, respectivamente.

Para a estimação dos parâmetros do modelo, novamente serão utilizadas duas abordagens: a clássica, utilizando a função **glm** do *software* R, e a Bayesiana, usando a função **stan_glm** do pacote **rstanarm** do *software* R, assumindo independência a priori entre os

coeficientes da regressão e distribuições a priori não informativas já fixadas por padrão na função **stan_glm**.

3.4.1 Estimação clássica

Os coeficientes estimados sob a abordagem clássica, com sexo feminino e com a faixa etária mais 65 anos como categorias de referência, bem como o p-valor do Teste de Wald para cada um dos parâmetros encontram-se na Tabela 12.

Com base no ajuste do modelo, observa-se que o p-valor de todas as covariáveis é menor do que o nível de significância de 5% definido, logo, utilizando o teste de Wald definido na Seção 2.3, há indícios de que existe associação significativa entre as variáveis explicativas e a variável resposta.

Tabela 12: Estimativas clássicas e p-valor do Teste de Wald para os parâmetros do Modelo de Regressão Poisson.

Covariável	Estimativa	Desvio Padrão	p-valor
Intercepto	0,26171	0,02559	$< 0,001$
Impressões	0,00534	0,00014	$< 0,001$
Custo	0,24844	0,00436	$< 0,001$
Sexo			
Masculino	-0,03344	0,01511	0,0269
Faixa Etária			
18 a 24	-0,35465	0,03979	$< 0,001$
25 a 34	0,18757	0,03060	$< 0,001$
35 a 44	0,17990	0,03124	$< 0,001$
45 a 54	0,25801	0,03079	$< 0,001$
55 a 64	0,32619	0,03004	$< 0,001$

Na Tabela 12, também podemos observar a faixa etária de 18 a 24 anos apresenta um efeito negativo, indicando uma redução na taxa de cliques em usuários que pertencem a esta faixa etária quando comparados à usuários na faixa etária acima dos 65 anos.

Chama atenção o fato de que a estimativa para o sexo masculino apresenta um efeito negativo na quantidade média de cliques, quando comparado ao sexo feminino; o que parece não estar de acordo com o observado na Seção 3.2, na qual observa-se que os usuários do sexo masculino são os que mais clicam nos anúncios, e com o resultado obtido

via modelo simples, apresentado na tabela 4, na qual a estimativa para o sexo masculino, considerando o sexo feminino como referência, foi positiva. Por outro lado, pode-se notar que o valor estimado é, em módulo, bem pequeno, indicando que o sexo não é uma característica tão relevante para explicar a taxa de cliques dadas as demais variáveis explicativas consideradas no modelo. Seguindo a interpretação dos resultados de acordo com a razão de taxas discutida na Seção 2.2.3, a Tabela 13 mostra os efeitos multiplicativos das variáveis explicativas na razão de taxas de cliques.

Tabela 13: Estimativas pontuais e intervalares (intervalos de confiança de 95%) para as Razões de taxas associadas os parâmetros do Modelo de Regressão Poisson.

	Razão de taxas	
Covariável	Estimativa pontual	Intervalo de credibilidade
Intercepto	1,29916	[1,23517 ; 1,36551]
Impressões	1,00535	[1,00508 ; 1,00562]
Custo	1,28203	[1,27108 ; 1,29301]
Sexo		
Masculino	0,96711	[0,93891 ; 0,99619]
Faixa Etária		
18 a 24	0,70142	[0,64863 ; 0,75812]
25 a 34	1,20632	[1,13628 ; 1,28112]
35 a 44	1,19709	[1,12620 ; 1,27292]
45 a 54	1,29435	[1,21877 ; 1,37515]
55 a 64	1,38568	[1,30671 ; 1,47001]

Analisando a estimativa para a covariável Impressões, para o aumento de uma unidade no número de impressões, espera-se o aumento de aproximadamente 0,5% na quantidade média de cliques dos usuários impactados pelo anúncio. Por outro lado, para a covariável Custo, com um aumento de R$1, espera-se o aumento de aproximadamente 28,2% na quantidade média de cliques dos usuários.

Conforme já discutido, percebemos a baixa influência do sexo na taxa de cliques dos usuários. De fato, de acordo com a Tabela 12, quando os clientes são do sexo masculino, espera-se uma redução de aproximadamente 3% na quantidade média de cliques em comparação com a categoria de referência sexo feminino.

Analisando a variável Faixa Etária, percebe-se os seguintes efeitos esperados em relação a categoria de referência (mais de 65 anos): uma redução de aproximadamente

30% na quantidade média de cliques para a faixa etária de 18 a 24 anos, um aumento de aproximadamente 20% para as faixas etárias de 25 a 34 anos e de 35 a 44 anos, um aumento de 29,4% para a faixa etária de 45 a 54 anos e um aumento de 38,5% para a faixa etária de 55 a 64 anos.

3.4.2 Estimação Bayesiana

Assim como na Seção 3.3.2, os resultados obtidos nesta seção foram baseados nas saídas da função **stan_glm**, assumindo 4 cadeias geradas, tendo cada uma 4.000 iterações, sendo descartadas as 2.000 primeiras como amostra de aquecimento (*burn-in*).

Inicialmente utilizou-se uma abordagem visual para avaliar a convergência, por meio da observação dos traços das cadeias dos parâmetros. Na Figura 4 (lado direito), pode-se observar os traços das cadeias dos coeficientes associados às variáveis Sexo Masculino e Custo, além da curva da densidade aproximada para a amostra obtida.

Ao verificar a Figura 4 para as variáveis Sexo Masculino e Custo, observar-se que a distribuição a posteriori é aproximadamente simétrica em ambos os casos. A posição central da distribuição indica o valor mais provável do parâmetro a posteriori, e é possível perceber que os traços das 4 cadeias são semelhantes e que as trajetórias estão estáveis, dando indícios de convergência para a distribuição a posteriori. O comportamento das cadeias dos demais parâmetros é análogo, podendo ser vistas no Apêndice 1. O diagnóstico de convergência formal pode ser obtido analizando a estatística *R-hat* de Gelman-Rubin apresentadas na Tabela 14.

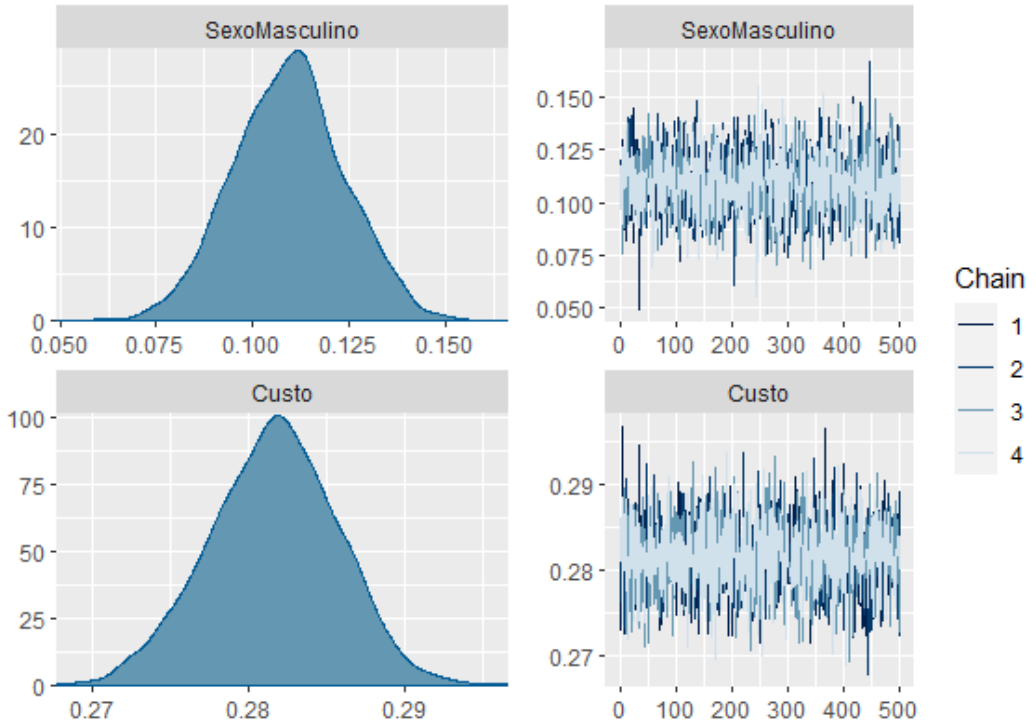

Figura 4: Densidade aproximada e Cadeias obtidas via HCM dos valores gerados para os coeficientes associados às variáveis explicativas Sexo Masculino e Custo.

A Tabela 14 também mostra a média a posteriori dos parâmetros, assim como as estimativas pontuais (que indicam os efeitos multiplicativos das variáveis explicativas na variável resposta) e os respectivos intervalos de credibilidade de 95% para as razões de taxas associadas a cada um dos parâmetros do modelo múltiplo.

Tabela 14: Estimativas pontuais para os parâmetros e estimativas pontuais e intervalares (intervalos de credibilidade de 95%) para as Razões de taxas e R-hat associadas os parâmetros do Modelo de Regressão Poisson Bayesiano.

	Coeficiente	Razão de taxas		
Covariável	Est. pontual	Estimativa pontual	Intervalo de credibilidade	R-hat
Intercepto	0,26122	1,29852	[1,23609 ; 1,36430]	1
Impressões	0,00534	1,00535	[1,00509 ; 1,00563]	1
Custo	0,24837	1,28193	[1,27141 ; 1,29294]	1
Sexo				
Masculino	-0,03363	0,96693	[0,94005 ; 0,99552]	1
Faixa Etária				
18 a 24	-0,35256	0,70289	[0,64894 ; 0,76076]	1
25 a 34	0,18816	1,20702	[1,13542 ; 1,28453]	1
35 a 44	0,17979	1,19697	[1,12701 ; 1,27310]	1
45 a 54	0,25866	1,29520	[1,21635 ; 1,37470]	1
55 a 64	0,32722	1,38710	[1,30714 ; 1,46986]	1

Nota-se que todas as estimativas (pontuais e intervalares) relacionadas às razões de taxas apresentadas nas Tabelas 13 e 14 são muito próximas, levando às mesmas interpretações feitas na Seção 3.4.1. Esta semelhança já era esperada, uma vez que os resultados obtidos pela abordagem Bayesiana basearam-se na utilização de distribuições a priori não informativas.

4 Conclusões

Esse trabalho iniciou observando o movimento dos empreendedores de tentar sair das dificuldades ocorridas nas empresas geradas pela COVID-19 e com a motivação de conseguir aumentar a performance dos anúncios para gerar mais retorno, verificando o que realmente está influenciando para ter mais cliques e quais características dos clientes influenciam na busca do produto.

O objetivo do trabalho foi tentar otimizar buscando uma melhora nas configurações dos anúncios, identificando as características dos clientes que mais buscam o produto da empresa e consequentemente geram mais cliques.

Observando a Figura 3, nota-se que o perfil de usuários que mais clicam nos anúncios são do sexo masculino na faixa etária de 35 à 54 anos.

Através das Tabelas 12 e 13, percebe-se que a característica que mais influencia negativamente na quantidade de cliques é ser da faixa etária de 18 a 24 anos. Estas estimativas dão indícios de que os anúncios não estão alcançando esse público; logo esse público poderia ser estudado e motivado com mais investimento para tentar atrair esses clientes para a empresa.

As conclusões acima mostram que esta modelagem oferece para os empreendedores uma oportunidade de entender e configurar melhor os seus anúncios, entendendo quais características dos usuários que mais buscam o seu produto, gerando uma otimização nas suas campanhas e convertendo mais clientes.

Vale destacar, que os dados utilizados nesse trabalho possuem algumas limitações, implicando em menos detalhes nos resultados. Outras variáveis importantes para uma melhor configuração de um anúncio feito na Ferramenta Google Ads devem ser investigadas. Por exemplo, informações de localidade e renda dos usuários são apresentadas na ferramenta, porém não foram utilizadas neste trabalho por estarem agregadas por questões de proteção dos dados dos usuários. Essas informações poderiam tornar as análises mais interessantes e geraria um retorno melhor para os usuários da ferramente Google Ads.

Vale considerar que foi utilizado o modelo generalizado da família Poisson e observa-se a média da variável resposta clique é de $3,8$ e que não é igual a variância que tem o valor de $10,8$, mostrando uma superdispersão nos dados. Portando, para trabalhos futuros, recomenda-se considerar a modelagem dos dados utilizando um modelo de superdispersão ou um modelo Binomial Negativo. Notou-se também que a estimativa dos parâmetros associados ao sexo masculino obtida nos modelos múltiplos não seguiu o padrão esperado com base na análise descritiva. Apesar do valor pequeno desta estimativa, indicando uma baixa influência da variável sexo nas taxas de cliques ao se considerar as demais variáveis explicativas, esse comportamento diferente do esperado levanta a suspeita de que pode haver colinearidade no modelo ou de que a influência de outros fatores estejam afetando o efeito desta covariável, sendo necessária uma investigação mais aprofundada no modelo.

Os resultados obtidos pela estimação clássica e pela estimação Bayesiana são bem próximos devido ao uso de uma priori não informativa. É interessante para trabalhos futuros analisar distribuições a priori que sejam informativas para a utilização do modelo com o ponto de vista Bayesiano como um modelo alternativo.

Referências

DOBSON, A. J.; BARNETT, A. G. *An Introduction to Generalized Linear Models*. [S.l.]: CRC Press, 2018. (Chapman & Hall/CRC Texts in Statistical Science).

EHLERS, R. S. *Estatística Computacional*. 2003.

GAMERMAN, D.; LOPES, H. F. *Markov chain Monte Carlo: stochastic simulation for Bayesian inference*. [S.l.]: CRC press, 2006.

GELMAN, A.; GILKS, W. R.; ROBERTS, G. O. Weak convergence and optimal scaling of random walk metropolis algorithms. *The annals of applied probability*, Institute of Mathematical Statistics, v. 7, n. 1, p. 110–120, 1997.

GELMAN, A.; RUBIN, D. B. Inference from Iterative Simulation Using Multiple Sequences. *Statistical Science*, Institute of Mathematical Statistics, v. 7, n. 4, p. 457 – 472, 1992.

GOODRICH, B. et al. *rstanarm: Bayesian applied regression modeling via Stan.* 2022. R package version 2.21.3. Disponível em: ⟨https://mc-stan.org/rstanarm/⟩.

GOOGLE. *Google Ads*. 2023. Acessado em 22/06/2023. Disponível em: ⟨https://ads.google.com/intl/pt-BR_br/home/⟩.

HABTE, A.; DESSU, S. The uptake of key elements of sexual and reproductive health services and its predictors among rural adolescents in southern ethiopia, 2020: application of a poisson regression analysis. *Reproductive Health*, BioMed Central, v. 20, n. 1, p. 1–13, 2023.

JORGENSEN, B. *Theory of Linear Models*. [S.l.]: CRC Press, 2019.

KUMAR, R. et al. Predicting clicks: CTR estimation of advertisements using logistic regression classifier. In: IEEE. *2015 IEEE international advance computing conference (IACC)*. [S.l.], 2015. p. 1134–1138.

MARTINS, F. *Otimização de uma campanha publicitária na rede de pesquisa do Google Ads utilizando Teoria da Decisão Bayesiana*. Tese (Doutorado) — Universidade de São Paulo, 2019.

METROPOLIS, N. et al. Equation of state calculations by fast computing machines. *The Journal of Chemical Physics*, American Institute of Physics, v. 21, n. 6, p. 1087–1092, 1953.

MIGON, H. S.; GAMERMAN, D.; LOUZADA, F. *Statistical Inference: An Integrated Approach*. [S.l.]: CRC Press, 2014. (Chapman & Hall/CRC Texts in Statistical Science).

NELDER, J. A.; WEDDERBURN, R. W. M. Generalized linear model. *Journal of the Royal Statistical Society*, v. 135, n. 3, p. 370–384, 1972.

NOGUEIRA, D. A. *Proposta e avaliação de critérios de convergência para o método de Monte Carlo via Cadeias de Markov: casos uni e multivariados*. Dissertação (Mestrado) — Universidade Federal de Lavras, 2004.

PéREZ, F. L. *Monte Carlo Hamiltoniano na regressão logística*. 2022. ⟨http://leg.ufpr.br/~lucambio/CE225/20212S/MLG_Monte_Carlo.html⟩. 2022.

R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria, 2022. Disponível em: ⟨https://www.R-project.org/⟩.

RAMOS, F. M. D. G. *Aplicação dos modelos lineares generalizados à previsão de reservas para sinistros*. Tese (Doutorado) — Instituto Superior de Economia e Gestão, 2000.

SEBRAE. *O Impacto da pandemia de coronavírus nos pequenos negócios*. 2021. ⟨https://fgvprojetos.fgv.br/sites/fgvprojetos.fgv.br/files/impacto-coronavirus-13aedicao_diretoria-v7.pdf⟩. 2021.

TOHARUDIN, T. et al. Bayesian Poisson model for covid-19 in West Java Indonesia. *Sylwan*, v. 164, n. 6, p. 279–290, 2020.

APÊNDICE 1 – Resultados obtidos via Monte Carlo Hamiltoniano para o Modelo Múltiplo

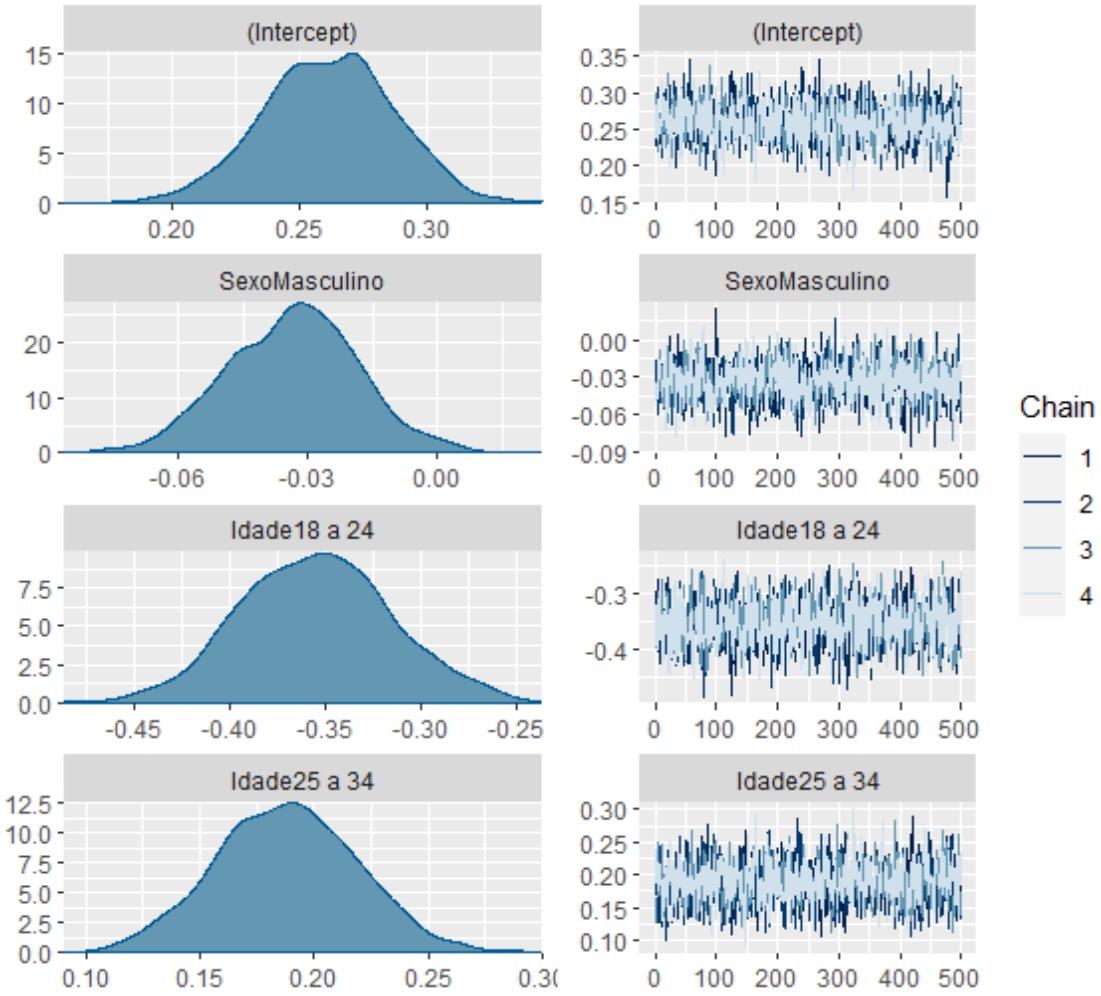

Figura 5: Densidade aproximada e Cadeias obtidas via HCM dos valores gerados para os coeficientes associados às variáveis explicativas e intercepto.

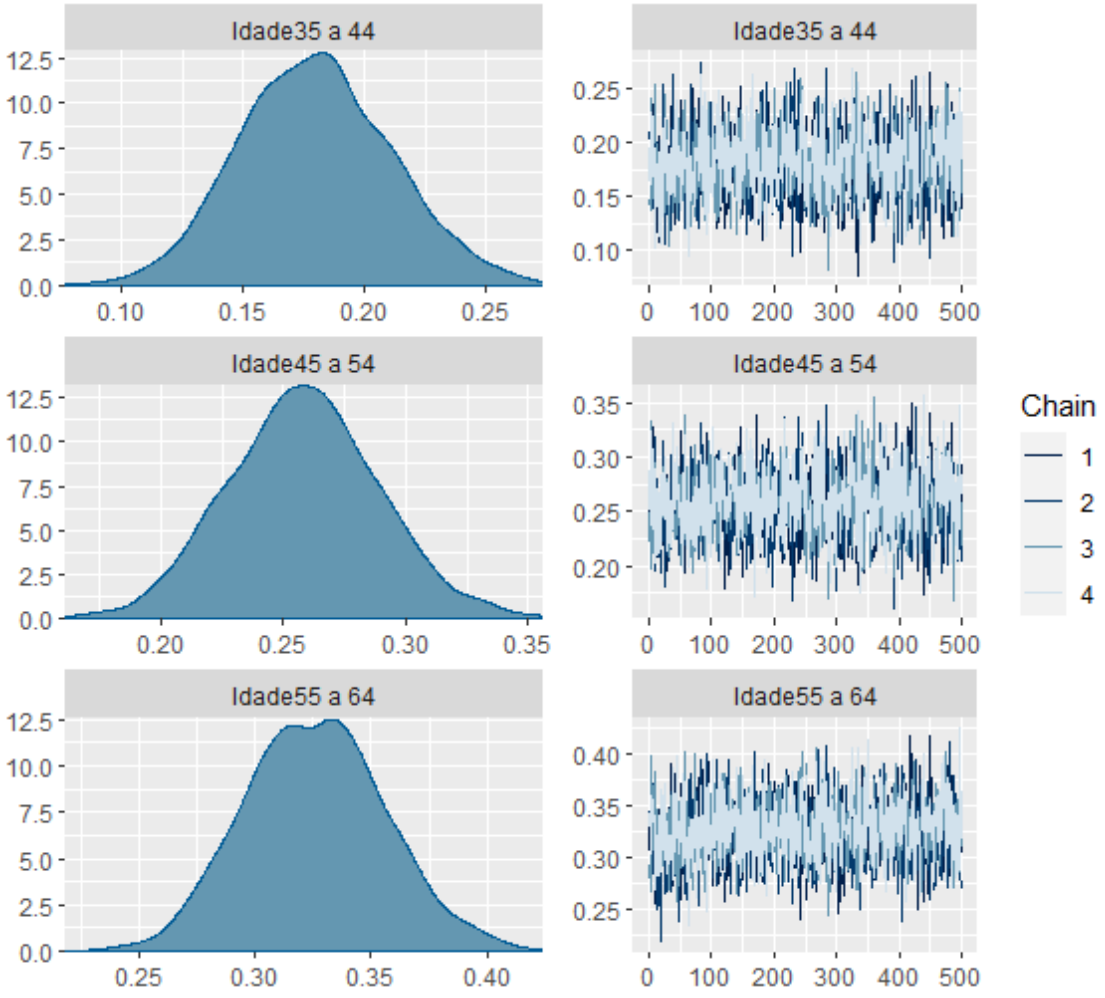

Figura 6: Densidade aproximada e Cadeias obtidas via HCM dos valores gerados para os coeficientes associados às variáveis explicativas (continuação).