UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE CIÊNCIAS EXATAS E DA NATUREZA
PROGRAMA DE PÓS-GRADUAÇÃO EM ESTATÍSTICA

CARLOS EDUARDO FRANTZ MANCHINI

PRIVACIDADE DIFERENCIAL: aplicação em modelos de regressão

Recife
2021

CARLOS EDUARDO FRANTZ MANCHINI

PRIVACIDADE DIFERENCIAL: aplicação em modelos de regressão

Dissertação apresentada ao Programa de Pós-Graduação em Estatística do Centro de Ciências Exatas e da Natureza da Universidade Federal de Pernambuco, como requisito parcial à obtenção do título de mestre em Estatística.

Área de Concentração: Estatística Aplicada

Orientador: Dr. Raydonal Ospina Martínez

Recife

2021

Catalogação na fonte
Bibliotecária Monick Raquel Silvestre da S. Portes, CRB4-1217

M268p Manchini, Carlos Eduardo Frantz
Privacidade diferencial: aplicação em modelos de regressão / Carlos Eduardo Frantz Manchini. – 2021.
81 f.: il., fig., tab.

Orientador: Raydonal Ospina Martínez.
Dissertação (Mestrado) – Universidade Federal de Pernambuco. CCEN, Estatística, Recife, 2021.
Inclui referências.

1. Estatística aplicada. 2. Regressão. I. Martínez, Raydonal Ospina (orientador). II. Título.

310 CDD (23. ed.) UFPE - CCEN 2021 – 33

CARLOS EDUARDO FRANTZ MANCHINI

PRIVACIDADE DIFERENCIAL: APLICAÇÃO EM MODELOS DE REGRESSÃO

Dissertação apresentada ao Programa de Pós-Graduação em Estatística do Centro de Ciências Exatas e da Natureza da Universidade Federal de Pernambuco, como requisito parcial à obtenção do título de mestre em Estatística. Área de Concentração: Estatística Aplicada

Aprovada em: 03 de fevereiro de 2021

BANCA EXAMINADORA

Prof. Dr. Raydonal Ospina Martínez (Orientador)
Universidade Federal de Pernambuco - UFPE

Prof. Dr. Francisco Cribari-Neto
Universidade Federal de Pernambuco - UFPE

Prof. Dr. Eufrásio de Andrade Lima Neto
Universidade Federal da Paraíba - UFPB

Com gratidão, dedico este trabalho aos meus pais, João e Helena, meus maiores exemplos.

AGRADECIMENTOS

Primordialmente, agradeço aos meus pais pelo apoio constante e incondicional. Por acreditarem no meu potencial e pela confiança a mim depositada para conclusão deste curso de pós-graduação.

Aos professores e mentores que fizeram parte do meu crescimento, agradeço pelo ensino de qualidade que proporciona uma formação com excelência a seus alunos, em especial ao Prof. Dr. Fábio Bayer, um exemplo de profissional ético e comprometido que incentivou a minha formação acadêmica.

Agradeço aos amigos de longa data que me sustentaram a manter-me firme nos estudos durante essa etapa. Sua ajuda foi fundamental para minha resiliência e sou imensamente grato pelos anos de amizade e parceria.

A toda minha família pelo suporte e pelos valores transmitidos que servem como guia para vivência em sociedade como um ser humano de caráter digno.

Ao meu orientador Prof. Dr. Raydonal Ospina, agradeço pela disponibilidade e dedicação, pela confiança a mim imposta e pelos ensinamentos – incluso os não-acadêmicos – durante a realização deste trabalho.

A Raquel, por toda compreensão, apoio e motivação. Por me fazer acreditar em mim mesmo, por fazer parte e me lembrar que eu tinha uma vida além dessa dissertação.

Compactuando com a campanha "*Thank you Edward Snowden*", agradeço a Snowden por abrir mão da sua liberdade americana ao denunciar e expor a indevida vigilância em massa regida pelo governo estadunidense. Considero-o um grande disseminador da politização a violação de privacidade na internet.

Um agradecimento especial à CAPES pelo auxílio financeiro fornecido durante todo o período de estudo.

A todos que de alguma forma contribuíram para conclusão do mestrado, direta ou indiretamente, muito obrigado!

RESUMO

A produção massiva de dados nesta era digital propicia violações à privacidade individual. A busca incessante por dados pessoais, torna cada vez mais recorrente a exposição dessas informações. O presente trabalho apresenta a área de privacidade diferencial, a qual fornece fortes garantias de confidencialidade aos dados e robustez contra ataques de re-identificação invasivos. A definição destaca-se na literatura por seu embasamento matemático rigoroso capaz de quantificar a perda de privacidade. A intenção é tornar os indivíduos indistinguíveis inserindo aleatoriedade no processo de estimação enquanto mantém a representatividade no estudo. Técnicas antecedentes para proteção de dados são revisadas (expondo vulnerabilidades) e diversas aplicações práticas são relatadas. Foi implementada uma abordagem diferencialmente privada dos modelos de regressão linear e logístico. No caso linear, considerou-se a presença de heteroscedasticidade, assim como estimadores consistentes para estimação adequada. A influência da restrição de privacidade no desempenho dos estimadores foi avaliada por meio de simulações de Monte Carlo, incluso as taxas de rejeição dos testes sob hipótese nula e alternativa. Os resultados da avaliação numérica evidenciaram maiores distorções inferenciais para restrições mais rigorosas de privacidade. Ao final, aplicações a dados reais são apresentadas e discutidas. Adicionalmente, o trabalho fomenta a reflexão do impacto tecnológico na contemporaneidade, tratando da posse e coleta de dados não-consentida relacionada ao capitalismo da vigilância e manipulação comportamental.

Palavras-chave: Anonimato. Exposição de dados. Informação de identificação pessoal. Privacidade diferencial. Regressão.

ABSTRACT

The massive data generation in the digital age leads to violations of individual privacy. The constant search for personal data becomes increasingly recurrent exposure of such information. This work presents the differential privacy research area, which provides strong guarantees of data confidentiality and robustness against invasive re-identification attacks. The definition excels in the literature for its rigorous mathematical background able to quantify the loss of privacy. The aim is to make individuals indistinguishable by adding randomness into the estimation process while maintaining statistical representativeness in the study. Previous data protection techniques are reviewed (exposing weakness) and many practical applications are reported. A differentially private approach of linear and logistic regression models was developed. In the linear model, we consider the presence of heteroskedasticity and consistent estimators for adequate estimation. The influence of privacy restriction on the performance of the estimators was evaluated using Monte Carlo simulations, including test rejection rates under null and alternative hypothesis. The results of numerical evaluation showed greater inferential distortions for stricter privacy restrictions. In the end, applications to real data are presented and discussed. Additionally, the work encourages reflection on the current technological influence, considering the possession and collection of data without consent associated with surveillance capitalism and behavioral manipulation.

Keywords: Anonymity. Data breach. Differential privacy. Personally identifiable information. Regression.

LISTA DE ABREVIATURAS E SIGLAS

DP *Differential Privacy* (Privacidade Diferencial)

DPBA *Differentiated Privacy Budget Allocation* (Alocação de privacidade diferencial)

FM *Functional Mechanism* (Mecanismo Funcional)

GDPR *General Data Protection Regulation* (Regulamento Geral de Proteção de Dados)

IBM *International Business Machines Corporation*

IoT *Internet of Things* (Internet das coisas)

IVA *Intelligent Virtual Assistant* (Assistente Virtual Inteligente)

LGPD Lei Geral de Proteção de Dados Pessoais

MLG Modelo Linear Generalizado

MVDP Máxima Verossimilhança Diferencialmente Privada

MW *Multiplicative Weights* (pesos multiplicativos)

PII *Personally Identifiable Information* (Informação de Identificação Pessoal)

SAP *Systeme, Anwendungen und Produkte in der datenverarbeitung*

SDC *Statistical Disclosure Control* (Controle Estatístico de Divulgação)

SIG Sistemas de Informação Geográfica

SQL *Structured Query Language* (Linguagem de consulta estruturada)

SUMÁRIO

1 INTRODUÇÃO

Atualmente, a sociedade vivencia uma crescente urgência em se manter conectada. De acordo com pesquisa, há mais de dois dispositivos digitais por habitante no Brasil, totalizando 424 milhões de aparelhos ativos (MEIRELLES, FGV, 2020). O desenvolvimento tecnológico impulsiona a disseminação da internet a nível global. Nos últimos 15 anos, a proporção de pessoas com acesso à rede *web* cresceu aproximadamente 330%, alcançando a marca de 4.8 bilhões de usuários (GROUP, 2020). O continente europeu e a América do Norte detêm as maiores proporções populacionais de utilização da internet ($\approx$ 90%). Em países desenvolvidos, essa porcentagem foi estimada em 86.6% no ano de 2019 (ITU, 2019). Apesar da pequena quantidade de indivíduos conectados em países subdesenvolvidos, o relatório das Nações Unidas (UNITED NATIONS, 2020) destaca um crescimento digital devido a expansão da cobertura de sinal 3G, acrescida de 51% em 2015 para 79% em 2019.

Grande parte das nossas atividades *online* envolvendo um aparelho conectado à internet geram rastros digitais em forma de dados. Essas informações pessoais estão sendo coletadas em tempo real sob diferentes plataformas e sendo utilizadas, por exemplo, para interesses econômicos (ZUBOFF, 2015). Diversos dispositivos eletrônicos têm capacidade de extrair dados relacionados a nossa identidade, desde smartphones até relógios de pulso inteligentes. A recente tecnologia IoT (*Internet of Things*) ameaça a violação de privacidade, pois sincroniza a internet com sistemas digitais de uso cotidiano como eletrodomésticos, automóveis, termostatos, entre outros (ACAR *et al.*, 2020). O projeto Haystack[1], iniciado em 2015, desenvolveu um aplicativo que analisa o tráfego móvel de dispositivos e identifica vazamentos de privacidade causados por outros aplicativos Android. Os pesquisadores do projeto descobriram que mais de 70% dos aplicativos conectava-se a rastreadores digitais que compartilhavam os dados à serviços terceiros.

Com ou sem consentimento, nosso comportamento está sendo monitorado e constantemente antecipado pela indústria que constrói seu modelo de negócios em torno do acúmulo e processamento de dados (RODOTÀ, 2015; HOOFNAGLE *et al.*, 2012). Corporações têm investido fortemente na busca de informações sobre seus usuários para investigação de hábitos e direcionamento segmentado de publicidade que estimule o consumo e gere lucro[2]. Não é coincidência que diariamente somos bombardeados por anúncios publicitários baseados em pre-

[1] Para mais informações do projeto, ver <https://haystack.mobi/>. Ferramenta interativa disponível em <https://haystack.mobi/panopticon>.

[2] O relatório divulgado pela PwC aponta que a publicidade digital ultrapassou a televisiva nos EUA totalizando uma receita de US$ 124.6 bilhões em 2019**.** Disponível em <https://www.iab.com/wp-content/uploads/2020/05/FY19-IAB-Internet-Ad-Revenue-Report_Final.pdf>.

ferências individuais. A globalização tecnológica e o desenvolvimento do capitalismo tornaram a plataforma eletrônica dominante na gestão de negócios, coletando e comercializando dados (WANG; LI, 2017).

Os avanços na infraestrutura e computação facilitaram a coleta, armazenamento e análise de grandes quantidades de dados. Com isso, a produção de dados atinge níveis gigantescos[3] e desencadeia desafios adicionais para a proteção da privacidade. Conforme estes dados tornam-se mais acessíveis e detalhados faz-se necessária a utilização de mecanismos que garantam o anonimato dos usuários ao serem disponibilizados no universo digital. O compartilhamento de dados sem proteção adequada pode acarretar, perante a participantes mal-intencionados em fraudes, extorsões e até falsidade ideológica (WINDER, 2019).

A violação de dados está se tornando cada vez mais recorrente. Apenas no segundo semestre de 2019, mais de 3800 ocorrências invasivas foram detectadas pela equipe de segurança cibernética RBS - Risk Based Security (2019) expondo 4.1 bilhões de registros pessoais. Em geral, esses dados provêm do setor corporativo e contêm informações pessoais e financeiras dos titulares (nome, *e-mail*, senha, renda, cartão bancário, etc.). O relatório atualizado da RBS para 2020 indicou uma redução no número de violações comparativamente ao ano anterior, contudo, bateu recorde na quantidade de dados expostos (27 bilhões). A vice-presidente da equipe problematiza os bancos de dados hospedados em sistemas mal-configurados como principal causa desse aumento e ressalta que a pandemia de coronavírus criou um ambiente mais suscetível a erros empresariais, os quais favoreceram a ação de criminosos cibernéticos (Risk Based Security, 2020).

Como constatado no estudo, há muitas vulnerabilidades no que diz respeito à segurança de dados. O abuso de dados é perigoso e tem capacidade de prejudicar, inclusive a democracia, influenciando acontecimentos importantes como as eleições norte-americanas de 2016, a qual originou o caso Facebook-Cambridge Analytica[4] (DELMAZO; VALENTE, 2018). Este visava persuadir a opinião de eleitores estadunidenses em prol de alguns candidatos e partidos políticos. Outra sinalização de fraude ocorre quando há falta de transparência por parte das empresas nos termos e condições de uso de uma determinada tecnologia. Muitas vezes as especificações são genéricas e "ocultam" as verdadeiras permissões de acesso a dados ou até mesmo quando o usuário por imprudência não verifica a política de privacidade e oportuniza a

[3] A pesquisa de 2017 estima que 2.5 quintilhões *bytes* de dados foram gerados por dia durante o ano. Ainda, declara que 90% da quantidade total de dados foi produzida exclusivamente nos dois últimos anos antecedentes à pesquisa. Acesso em <https://www.domo.com/learn/data-never-sleeps-5>.

[4] Detalhado na Seção 2.5.1.

isenção de empresas mal-intencionadas (OBAR; OELDORF-HIRSCH, 2020).

Preservar a confidencialidade de dados é um campo de estudo em ascensão que acompanha a evolução tecnológica e renova-se com o passar do tempo. Em contradição, esses avanços também favoreceram o desenvolvimento de técnicas para reconstruir bancos de dados sigilosos (DWORK; MCSHERRY; TALWAR, 2007). À primeira vista, pode parecer que uma simples técnica de anonimização, como a remoção de nomes e identificadores, aplicada aos dados é suficiente para preservar a privacidade. Ainda assim, é possível capturar traços para identificar indivíduos indiretamente e violar sua proteção (ROCHER; HENDRICKX; MONTJOYE, 2019). Logo, recomenda-se adotar uma abordagem para preservação de privacidade com embasamento matemático rigoroso que forneça modelos seguros e flexíveis que se adequem a diversos conjuntos de dados e permitam a extração de conhecimento útil (JAIN; GYANCHANDANI; KHARE, 2016).

Há diversas áreas acadêmicas (majoritariamente em ciência da computação) debatendo sobre privacidade de dados, adaptando-se aos seus contextos e propondo métodos que forneçam proteção para resolução de problemas específicos (FUNG *et al.*, 2010). A predominância computacional na área carece de embasamento inferencial, deste modo, necessita-se a inserção de profissionais devidamente qualificados em estatística, ciência de dados e afins para o aprimoramento das pesquisas nessa área. O principal desafio consiste em projetar sistemas e técnicas de processamento para fazer inferência sobre dados sensíveis, mantendo a privacidade e segurança das identidades individuais.

A área de Privacidade Diferencial (*Differential Privacy* - DP) (DWORK *et al.*, 2006) ganhou destaque no desenvolvimento e implementação de métodos privados demandada pela crescente exposição de dados pessoais. As garantias de privacidade estabelecidas em DP apresentam rigor matemático e são de natureza estatística, diferente das garantias empregadas no k-anonimato (SWEENEY, 2002) e teoria da informação (SANKAR; RAJAGOPALAN; POOR, 2013).

Muitas tarefas analíticas de informação contendo o processamento de dados estão sendo adaptadas para obtenção de resultados úteis que preservem a privacidade dos indivíduos. Esta modificação foi exigida, principalmente, por ataques de inversão que permitiram revelar informações pessoais confidenciais (FREDRIKSON; JHA; RISTENPART, 2015). No âmbito do aprendizado de máquina, denomina-se "inversão de modelo" quando um invasor, por meio das estimativas públicas resultantes do ajuste, re-identifica as observações individuais (SHOKRI *et al.*, 2017). Assim, os dados sensíveis que incorporam métodos usuais estão sob ameaça e

vulneráveis a exposições, especialmente na esfera médica, financeira e outras altamente sigilosas. O estudo de caso de Fredrikson *et al.* (2014) apresenta um classificador linear capaz de prever o genótipo dos indivíduos (utilizados como covariáveis na modelagem) via ataque de inversão do modelo.

Em modelos de regressão, a maioria das soluções existentes contra ataques de inversão são pouco satisfatórias, pois produzem estimativas fortemente viesadas devido à grande quantidade de ruído imposto durante o processo de estimação para alcançar o anonimato (WANG; SI; WU, 2015). A contaminação excessiva deteriora a capacidade preditiva do modelo e, consequentemente, compromete a acurácia de estimativas representativas. Outra limitação encontra-se nas propostas diferencialmente privadas restritas a classes de regressão pouco usuais, as quais dificultam a acessibilidade e reprodutibilidade científica. A ideia central da análise de regressão combinada à privacidade diferencial é adicionar ruído controlável em uma determinada etapa da modelagem, garantindo equilíbrio entre preservação da privacidade e utilidade (DWORK; ROTH, 2014).

1.1 OBJETIVO GERAL

O presente trabalho propõe uma abordagem diferencialmente privada dos modelos de regressão linear e logístico por meio da perturbação.

1.2 OBJETIVOS ESPECÍFICOS

Implementar os modelos de regressão linear e logístico que satisfazem as garantias diferencialmente privadas através da perturbação objetiva.

Considerar uma estrutura de heteroscedasticidade na modelagem de regressão linear DP, que a melhor de nosso conhecimento, é inexistente na literatura.

Avaliar numericamente o desempenho dos modelos propostos sob diferentes graus de privacidade através de simulações de Monte Carlo e compará-los com alternativas da literatura DP em aplicações com bancos de dados reais.

Realizar uma revisão bibliográfica de métodos para proteção de dados incluindo DP e suas aplicações práticas.

Sintetizar um estudo das diferentes concepções de privacidade associado a posse de dados na contemporaneidade.

Destaca-se a relevância desse trabalho no âmbito da pesquisa científica brasileira,

embora DP tenha um grande potencial de crescimento fomentado pelas novas tendências tecnológicas, ainda há poucos trabalhos publicados nessa área no Brasil.

2 PRIVACIDADE: CONCEITOS E DISCUSSÕES

> "*Uma transformação muito mais complexa e interessante está apenas começando, movida por uma crescente tensão entre duas forças distintas: o velho e o novo poder*" (HEIMANS; TIMMS, 2014).

2.1 (RE)DEFINIÇÕES CONCEITUAIS

A compreensão do conceito de privacidade entre os indivíduos, além de ser subjetiva, varia no decorrer da história. Conceituações heterogêneas podem ser encontradas em diversos campos científicos, jurídicos e governamentais. De modo geral, a busca por privacidade é essencial para garantia da democracia e do direito à livre escolha (SOLOVE, 2008). Enquanto condição fundamental, as esferas estatais têm o dever de assegurar privacidade aos seus cidadãos. Ainda que não exista um conceito único e universal de privacidade, tal qual a concebemos hoje, tem sua primeira menção no século XIX (MACHADO, 2014).

Os advogados Warren e Brandeis formularam o conceito mais tradicional no Ocidente em 1890. O artigo de cunho jurídico "The Right to Privacy[1]" publicado na *Harvard Law Review*, define privacidade como *"the right to be let alone"* (direito de ser deixado só) (WARREN; BRANDEIS, 1890). Os autores defendem a privacidade como uma oposição aos regimes opressivos visionando uma reforma na lei em resposta à mudanças tecnológicas, prezando pelo direito de escolha e propriedade. Este conceito foi alvo de muitas reformulações, a exemplo, o apresentado pelo livro "Privacy and Freedom[2]" que consta: "é a reivindicação de indivíduos, grupos ou instituições, de que lhes seja permitido determinar quando, como e até que ponto as informações sobre si próprios sejam transmitidas a outros" (WESTIN, 1967, p. 7).

Na década de 1940, as Nações Unidas pronunciou em uma declaração universal de direitos humanos, a garantia de privacidade como uma responsabilidade civil, *in verbis*:

> Art. 12º *"No one shall be subjected to arbitrary interference with his privacy, family, home or correspondence, nor to attacks upon his honor and reputation. Everyone has the right to the protection of the law against such interference or attacks[3]"* (United Nations General Assembly, 1948).

Os primeiros termos relacionados ao atual conceito de privacidade no Brasil aparecem na Constituição Federal de 1988, através dos vocábulos "intimidade" e "vida privada" (MACHADO,

---

1 O Direito à Privacidade.

2 Privacidade e Liberdade.

3 Tradução livre: "Ninguém será sujeito a interferências em sua vida privada, na sua família, em seu domicílio ou na sua correspondência, nem a ataques à sua honra e reputação. Todo ser humano tem direito à proteção da lei contra tais interferências ou ataques".

2014). Ainda que não associadas diretamente ao conceito moderno de privacidade, fazem alusão à noção de direito de *escolha* e de *estar* só, apresentado por Brandeis e Warren (RODOTÀ, 2015).

A privacidade enquanto conceito contemporâneo surge, precipuamente, com o advento da tecnologia e das redes sociais; o privado molda-se pela escolha de expor/omitir informações pessoais (RODOTÀ, 2015). As múltiplas possibilidades de divulgação da vida privada no universo digital acarretam em uma produção massiva de dados pessoais. O valor destes é traduzido em transações comerciais, muitas vezes mediadas por exigências governamentais, que acabam tornando os dados "públicos" passíveis de manipulação. Desta forma, o conceito aplicado nesta pesquisa baseia-se em preceitos contemporâneos sobre o direito à privacidade, sendo este "caracterizado pela liberdade de autodeterminação informativa, isto é, a capacidade de controlar as informações pessoais pelo seu titular" (MACHADO, 2014, p. 339).

Em consequência à crescente disponibilidade de dados, as legislações mundiais viram-se forçadas a se adaptarem para proteger a privacidade dos indivíduos. Atualmente, mais de 130 países adotam estruturas legislativas direcionadas à proteção de dados mantidos por órgãos públicos e privados (GREENLEAF, 2019). Na Europa, está vigente desde 2018 o Regulamento Geral de Proteção de Dados (GDPR) que obriga toda empresa que colete dados pessoais a obter o consentimento prévio dos usuários[4]. Inspirados pela GDPR, no Brasil, foi aprovada e sancionada em 2018, entrando em vigor em setembro de 2020, a Lei Geral de Proteção de Dados (LGPD) nº 13.709/18 que regulamenta a política de proteção e privacidade de dados pessoais no quesito coleta, armazenamento e compartilhamento. A lei empodera os titulares permitindo-os retificar, cancelar e solicitar a exclusão de suas informações da base de dados central. Empresas, órgãos públicos e *websites* possuem o prazo máximo de um ano passível de punição para alterar drasticamente suas operações com tratamento de dados e se adequarem às novas políticas. Cabe salientar que a LGPD, juntamente aos comitês de ética, não estabelecem restrições legais ao uso de dados pessoais em situações de interesse público.

Nos Estados Unidos, uma norma severa é imposta no Departamento do Censo (Census Bureau), a qual está codificada no código americano (*Title 13 – U.S. Code*). Em suma, o título declara: É ilegal qualquer publicação de dados privados fornecidos por estabelecimentos ou indivíduos. A violação do compromisso de confidencialidade pode resultar em multa de até US$ 250.000 e sentença a prisão federal por até cinco anos. Por outro lado, o desperdício de

[4] Para mais detalhes do processo de formulação da GDPR e dos critérios regulatórios estabelecidos na União Européia, ver <https://ec.europa.eu/info/law/law-topic/data-protection_en>.

informação acontece quando estes dados acabam não sendo liberados por falta de mecanismos adequados que protejam a identidade pessoal da população, limitando, assim, o progresso científico e empreendedor.

2.2 PARADOXO DE PRIVACIDADE

Na contemporaneidade, onde dados pessoais são gerados ininterruptamente, a privacidade pode ser vista como paradoxal quando sua violação se dá pela vontade da própria vítima[5] (NORBERG; HORNE; HORNE, 2007). Embora o indivíduo defenda a proteção de suas informações, seu comportamento não reflete tal preocupação, favorecendo a divulgação das mesmas (TADDICKEN, 2014). Por exemplo, o usuário torna-se contribuinte quando concorda com os termos e condições das redes sociais, já que as informações coletadas a seu respeito passam a ser propriedade das empresas. Estas, são capazes de intrincar propositalmente as especificações (textos complexos, vagos, excessivamente longos e até letras miúdas) para omitir a autorização de acesso aos dados dos consumidores. As empresas ainda podem isentar-se da responsabilidade de confidencialidade acobertada por uma política de privacidade tendenciosa.

O estudo *"Privacy and Location Data Global Consumer Study"*, proposto pela empresa HERE Technologies (2018) via mapeamento de dados, foi realizado com oito mil pessoas em oito países[6]. Este revelou que 76% dos entrevistados sentem-se preocupados e vulneráveis sobre o modo como empresas coletam e utilizam seus dados de localização. No entanto, mais de dois terços dos consumidores sinalizou disposição a compartilhar seus dados digitalmente. Cerca de 65% disseram ter divulgado (com consentimento) ao menos uma vez suas informações de localização. Apesar das pessoas expressarem receio às práticas dos coletores sob seus dados, a maioria está propensa a compartilhá-los e não se dedica ativamente para protegê-los.

Curiosamente, o relatório (HERE Technologies, 2018) inferiu entre os cidadãos dos países estudados: os brasileiros são mais "entusiasmados" a compartilhar seus dados pois acreditam ser vital no mundo atual e confiam fielmente nos coletores e regulamentos jurídicos impostos. A França é o país em que há mais incoerência entre a omissão e divulgação de suas informações, ou seja, onde o paradoxo da privacidade é mais evidente. Os japoneses demonstram-se cautelosos e restringem fortemente o acesso aos dados, todavia, não hesitam em liberá-los para maior comodidade e economia de tempo.

---

[5] Compreende-se como vítima o usuário que fornece os dados e reivindica proteção.

[6] Incluso Alemanha, Austrália, Brasil, Estados Unidos, França, Holanda, Japão e Reino Unido.

Em 2019, a pesquisa foi atualizada considerando mais de dez mil consumidores e demonstrou maior consciência do público quanto ao compartilhamento de dados pessoais. Nove em cada dez consumidores disseram entender o valor das suas informações para a indústria e 75% ainda mostra-se preocupados em liberá-las aos serviços digitais. Contraditoriamente, 70% afirmam compartilhar sua localização ao menos ocasionalmente. Em particular no Brasil, essa taxa apresentou aumento de sete pontos percentuais em comparação ao estudo anterior (HERE Market Intelligence, 2019).

2.3 SOBERANIA DOS DADOS

Muitos estabelecimentos e governos totalitários apostam na imposição de vigilância constante como forma de estabelecer disciplina e controle. Pesquisas realizadas no campo da sociologia expõem o Panóptico como representação para este tipo de regime (BENTHAM, 2019). O Panóptico é uma estrutura arquitetada para cárceres e prisões em formato circular contornando uma torre central (Figura 1) projetada para ser vista como opaca e impossibilitar a identificação dos vigilantes pelos indivíduos permanentemente observados. Para o filósofo Michel Foucault, o panóptico objetiva induzir no detido um estado consciente e permanente de visibilidade que assegura o funcionamento autoritário e soberano do poder (FOUCAULT, 2012).

Figura 1 – Presídio Modelo em Cuba construído sob sistema panóptico.

Fonte: Hidden Architecture[7]

Analogicamente no mundo globalizado, a sociedade digital apresenta uma estrutura panóptica onde o controle é possível pela conectividade assídua ao invés do isolamento prisional. A dominância hierárquica promove o acesso exclusivo a certas informações que garantem a posse e uso de dados da forma que lhes convêm. Considerando que na democracia contemporânea há diversos sistemas orientados a dados (como vigilância, manipulação de informação, marketing direcionado, entre outros), o Estado como detentor das informações, e consequentemente de

[7] Disponível em: <http://hiddenarchitecture.net/panopticism-presidio-modelo>.

dados pessoais, possui maiores recursos para controle sobre a população[8]. Na época atual em que a Biopolítica Digital[9] predomina, uma nova forma de soberania surge, como reitera Byung-Chul Han, soberano agora é aquele que dispõe de dados (HAN, 2018).

Nesta era da informação, a fronteira entre privado e público torna-se cada vez mais tênue. Agamben (2009) reforça:

> "Não seria provavelmente errado definir a fase extrema do desenvolvimento capitalista que estamos vivendo uma gigantesca acumulação e proliferação dos dispositivos. Certamente, desde que apareceu o *homo sapiens* havia dispositivos, mas dir-se-ia que hoje não haveria um só instante na vida dos indivíduos que não seja modelado, contaminado ou controlado por algum dispositivo" (AGAMBEN, 2009, p. 42).

O filósofo considera dispositivo como qualquer coisa que tenha a capacidade de captar, orientar, determinar, interceptar, modelar, controlar e assegurar os gestos, as condutas, as opiniões e os discursos dos seres viventes. Neste contexto, dispositivo pode ser visto como a ferramenta responsável pela vigilância individual e coletiva.

O presente momento sofre muitas transformações capazes de induzir uma representação viesada da realidade. Grande parte do processo de formação da opinião é mediado pelo conteúdo a que o público é exposto, geralmente proveniente dos meios de comunicação. Com isso, uma polarização é potencializada pela tendenciosidade da mídia (PERSE; LAMBE, 2016). A "verdade" é ditada pelos resultados das ferramentas de busca, predominantemente via Google, o qual conhece mais cada indivíduo do que os mesmos (CADWALLADR, 2016). As interações nas redes sociais também podem ser direcionadas com o uso não-transparente de *big data* para fins de interesse próprio (comercial, político ou financeiro) (SCHÖNBERGER; CUKIER, 2013; CHEN; MAO; LIU, 2014). A dominação da mídia digital distorce a percepção humana e, consequentemente, modifica o comportamento (WAHLBERG; SJOBERG, 2000). Neste sentido, Gerbner *et al.* (1986) corrobora ao confirmar o impacto do conteúdo televisivo sobre o discernimento dos telespectadores. A situação ainda é agravada quando a manipulação é imperceptível, visto que dificulta compreender as consequências dessa influência na sociedade.

---

[8] Um escândalo envolvendo o governo ocorreu quando Edward Snowden denunciou a Agência de Segurança Nacional dos Estados Unidos por práticas ilegais de vigilância. Como ex-contratado, Snowden tornou público documentos confidenciais sobre espionagem clandestina da agência (CHEN, 2017).

[9] Conjunto de mecanismos e procedimentos tecnológicos (saber-poder) que tem como intuito manter e ampliar uma relação de domínio.

2.4 SETOR CORPORATIVO

Na era digital de *big data*, aprendizado de máquina e inteligência artificial, a coleta de informações pessoais e comportamentais é fortemente demandada para criação de softwares otimizados e modelos que representem uma simplificação da realidade. Dessa forma, a preocupação com a privacidade dos usuários aumenta. Organizações e agências estatísticas que trabalham com pesquisas e coleta de dados têm o dever de proteger a confidencialidade individual. Muitas delas têm investido em técnicas que permite o usuário realizar análises bloqueando o contato direto com os dados. Por exemplo, no acesso remoto a *query systems* (sistemas de consulta) realiza-se uma solicitação ao servidor que retorna apenas o resultado/modelo final. Esse sistema está sendo usado por diversas agências governamentais como Census Bureau (MACHANAVAJJHALA *et al.*, 2008) e Australian Bureau of Statistics (CHIPPERFIELD; GOW; LOONG, 2016). É imprescindível que os termos e condições na coleta e uso dos dados sejam explícitos e que mantenham a privacidade individual de cada consumidor.

A área de saúde pública requer atenção extra pois contém informações sensíveis dos pacientes e a utilização de dados pessoais é comumente liberada sob jurisdição sem o consentimento do paciente. No Brasil, sua permissão é concedida em circunstâncias emergenciais como segurança ou calamidade pública conforme estabelecido na LGPD. Por exemplo, a base de dados brasileira do Sistema Único de Saúde (SUS) constitui o Gerenciador de Ambiente Laboratorial (GAL) com objetivo de informatizar o sistema nacional de vigilância epidemiológica/ambiental, monitorar doenças, entre outros. Nos EUA, foi implementado um regulamento de utilidade pública, denominado HIPAA (*Health Insurance Portability and Accountability Act*), o qual detalha o procedimento de anonimização de dados em concordância com a lei estadunidense e canadense para proteger as informações dos pacientes (EDEMEKONG; ANNAMARAJU; HAYDEL, 2020).

O compartilhamento de dados é crucial para melhorar a qualidade dos serviços, reduzir os custos médicos e acelerar as pesquisas biomédicas (LO'AI *et al.*, 2016), essencialmente no atual período de coronavírus, e portanto deve ser realizado com os devidos protocolos de uso. As instituições de saúde usualmente adotam a metodologia de *de-identification* (desidentificação) nos dados para preservar a privacidade dos pacientes (GARFINKEL, 2015). Um exemplo sugestivo de simples implementação para evitar exposições de pacientes: suponha que se deseja determinar o número de pessoas diagnosticadas com determinada doença, então divulga-se os resultados de forma agregada. Assim, apenas inferências úteis acerca da população são extraídas

e a privacidade é preservada.

Contudo, muitos desses métodos apresentam proteção falha como mostrado em Emam *et al.* (2011) no âmbito da saúde pública. Um exemplo prático é apresentado em Woodward (1996). O autor descreve o caso de um banqueiro filiado a uma comissão estadual de saúde que teve acesso a uma lista de pacientes com câncer e os associou a clientes devedores do seu banco para lhes cobrar empréstimos pendentes. Outro exemplo de re-identificação aconteceu em Nova York, devido a políticas federais de acesso público a dados governamentais *(Freedom of Information Law*) onde foram divulgados os dados de 173 milhões de corridas de táxi durante 2013. As informações incluíam: horários com localização de partida e chegada, valor das tarifas e um identificador para cada táxi (a priori anonimizado). Posteriormente, descobriu-se que era possível vincular os passageiros aos seus endereços com as coordenadas precisas fornecidas pelo GPS (HERN, 2014).

Recentemente, a tecnologia *blockchain* tornou-se um recurso relevante para proteção e privacidade de dados (ZYSKIND; NATHAN; PENTLAND, 2015). Sua arquitetura descentralizada naturalmente fornece segurança contra ataques cibernéticos e viabiliza as garantias ao anonimato, pois − como afirma Revoredo (2019) − permite transações de dados e valor sem intermediários, com transparência e imutabilidade das transações gravadas em sua rede, cujos participantes chegam a um consenso sobre o verdadeiro estado dos registros compartilhados (REVOREDO, 2019, p. 233). Dessa forma, os usuários conseguem fazer transações sem revelar sua identidade através de chaves criptografadas, fato este impossibilitado no sistema econômico tradicional.

A literatura vem explorando aplicações médicas baseadas em *blockchain*. O trabalho de Yue *et al.* (2016) propõe um aplicativo arquitetado via *blockchain* para que o paciente possa controlar e processar seus próprios dados sem violar a privacidade. Dagher *et al.* (2018) estuda um equilíbrio entre privacidade e acessibilidade em registros de saúde com uma estrutura baseada em *blockchain* para fornecer proteção. Apesar do potencial promissor, há controvérsias relacionadas aos princípios de segurança do *blockchain*; a disposição aberta aos dados, transparência e imutabilidade, são aspectos avessos à privacidade "padrão" que infringem, por exemplo, o direito do cidadão de remover suas informações do banco de dados, visto que nenhuma transação pode ser revertida nem sequer alterar os dados (JOSHI; HAN; WANG, 2018).

2.5 COMPARTILHAMENTO DE DADOS

Os bancos de dados atuais gerados por diferentes fontes de informação, apresentam uma crescente disponibilidade de dados do tipo PII (informações de identificação pessoal), os quais favorecem a exposição e comprometem a privacidade. Sensores avançados de equipamentos tecnológicos (câmera, localizador GPS, microfone, acelerômetro, barômetro) possibilitam a digitalização massiva do comportamento humano. As funcionalidades destes dispositivos permitem armazenar dados pessoais como nome, sexo, idade, endereço, CEP, sintomas, medicações, localização, preferências, etc. A seguir, relatamos diversos casos em que esses dados foram usados indevidamente e beneficamente.

2.5.1 Utilização indevida

Nos últimos anos, muitos foram os escândalos relacionados ao abuso e vazamento de dados, inclusive em processos eleitorais. O escândalo do Facebook - Cambridge Analytica em 2016 foi um dos maiores casos de utilização indevida de informações pessoais para fins de publicidade política. A criação de um teste de personalidade no Facebook deu início à narrativa em que mais de 87 milhões de usuários participaram. Estes dados pessoais, coletados sem consentimento, foram vendidos à empresa Cambridge Analytica, a qual fazia parte do plano estratégico da campanha eleitoral de Donald Trump operando o marketing. Em seguida, foi desenvolvido um modelo complexo com aproximadamente 5 mil observações para cada eleitor. O intuito era inferir sobre a personalidade dos indivíduos para manipular os mais suscetíveis a anúncios ou *fake news*. Com esses perfis traçados, milhões de dólares foram gastos direcionando discursos customizados às pessoas no Facebook a favor de Trump. Após a descoberta reportada pelo jornal *The Guardian* (CADWALLADR, 2018) e com a contribuição de um ex-funcionário da Cambridge Analytica, o Facebook teve um prejuízo bilionário e reconheceu via Mark Zuckerberg (fundador e CEO) que a empresa não tomou os devidos cuidados na proteção dos dados utilizados por terceiros. Já a Cambridge Analytica que processou os dados teve falência decretada.

Os avanços no reconhecimento de fala e sua alta capacidade de processamento via redes neurais e computação paralela permitiram aplicações em dispositivos mais próximos dos usuários. A popularização dos assistentes virtuais inteligentes (IVA) ameaçam a privacidade, pois processam grandes quantidade de dados comportamentais com finalidade suspeita (CHUNG; LEE, 2018). Uma vez que boa parte dos *softwares* comerciais que compõem os aparelhos são restritos apenas ao desenvolvedor/proprietário (código-fonte fechado), torna-se difícil determinar

o emprego dos dados e a suscetibilidade à manipulações fraudulentas. Os dispositivos, por sua própria natureza, estão gravando e processando áudio ininterruptamente já que comandos específicos de voz são necessários para ativação das funções. Por exemplo, Chung *et al.* (2017) relata incidentes e destaca os riscos associados à utilização de IVAs.

Algumas polêmicas envolvendo assistentes de voz inteligentes do Google, Apple e Amazon ocorreram pela falta de transparência com os dados de seus usuários (ALEPIS; PATSAKIS, 2017; PFEIFLE, 2018). Essas "gigantes tecnológicas" empresas já tiveram o reconhecimento de fala suspenso em seus aparelhos ao menos uma vez historicamente por queixas ou processos judiciais. Segundo artigo do jornal *Irish Examiner*, funcionários terceirizados da Apple estavam analisando "em segundo plano" mais de mil áudios a cada turno registrados pela assistente Siri. Tal captura foi justificada ao refinamento da capacidade de interpretação da assistente. As gravações continham informações sensíveis e poderiam ser usadas indevidamente. Suspeita-se que após a denúncia mais de 300 pessoas foram demitidas (CASEY, 2019).

O mesmo ocorreu em 2015 com o reconhecimento de voz em *smart TVs* da empresa sul-coreana Samsung. A equipe assegurou proteção na transmissão das conversas gravadas pelo televisores ao centro de armazenamento e negaram qualquer possibilidade de hackeamento. Porém, foi mostrado que as informações nem sempre estavam criptografadas e poderiam ser expostas (KELION, 2015). Para mais referências relacionadas a exposição de dados em dispositivos domésticos digitais, ver Li *et al.* (2015b); Ren *et al.* (2019) e Acar *et al.* (2020).

2.5.2 Utilização benéfica

Conjuntamente às garantias de proteção a privacidade, o compartilhamento de dados pessoais pode ter sua utilização aliada à causas positivas como na área da saúde (PRICE; COHEN, 2019) ou segurança (HEATH, 2018). Pesquisadores de diversas áreas acadêmicas beneficiam-se de informações compartilhadas no mundo digital para estudo a favor da humanidade. Há projetos de pesquisas impactantes em curso que contribuem ao avanço da medicina moderna e biologia humana, como a identificação de variantes genéticas causadoras de doenças. Trata-se do *Personal Genome Project Canada* (REUTER *et al.*, 2018) e Kohane (2011). Com a aplicação de métodos diferencialmente privados, muitos dados que antes estavam retidos, agora podem contribuir para uma tomada de decisão mais assertiva sem violações de privacidade.

O projeto *500 cities*[10] fornece estimativas das 500 maiores cidades nos Estados Unidos contendo indicadores de comportamento não-saudável, doenças crônicas e cuidados

[10] Acesso em <https://www.urban.org/sites/default/files/publication/90376/500-cities-project_1.pdf>.

preventivos. Além da identificação de doenças emergentes, o estudo objetiva sustentar um planejamento mais eficaz em intervenções na saúde pública. Outro acontecimento foi a criação de um banco de dados no Reino Unido contendo informações de estudantes voltado para gestão e melhorias na medicina e educação (DOWELL *et al.*, 2018). A literatura médica estuda fortemente o reconhecimento de padrões para fazer previsões e antecipar prognósticos. Algoritmos de inteligência artificial têm sido capazes de identificar células cancerígenas na pele com mesma precisão que dermatologistas treinados (ESTEVA *et al.*, 2017).

Destaca-se a contribuição de dados sensíveis no combate à pandemia de coronavírus. Informações georreferenciadas capturadas de *smartphones* foram usadas para rastreamento e controle da população em locais públicos de aglomeração. Estes Sistemas de Informação Geográfica (SIG) são de grande utilidade contra a disseminação de doenças, pois auxiliam no monitoramento do contágio e na tomada de decisão das autoridades de saúde. Neste sentido, o Consórcio Europeu desenvolveu o PEPP-PT (*Pan-European Privacy-Preserving Proximity Tracing*)[11] vinculado ao Bluetooth dos celulares visando interromper novas cadeias de transmissão do COVID-19. O aplicativo notifica a proximidade com possíveis infectados para evitar contato e promete preservação total de privacidade aos usuários[12]. Epidemiologistas dizem que este tipo de rastreamento será vital para conter o crescimento exponencial de propagações em epidemias futuras. Ainda, com o progresso de compartilhamento seguro de dados, salienta-se o potencial de DP na área de processamento de imagens para identificação de pessoas (desaparecidas/foragidas) via reconhecimento biométrico e facial (CHAMIKARA *et al.*, 2020).

[11] Para mais detalhes: <https://www.pepp-pt.org>.

[12] Alguns projetos e instituições que monitoram o avanço da pandemia foram contestados por utilizar dados pessoais e expor os pacientes. A exemplo de três acontecidos noticiados: <https://olhardigital.com.br/2020/12/31/seguranca/covid-19-grupo-nso-usou-dados-privados-reais-para-oferecer-software-de-rastreamento/>; <https://www.hrw.org/news/2020/12/15/personal-data-thousands-covid-19-patients-leaked-moscow/> e <https://saude.estadao.com.br/noticias/geral,vazamento-de-senha-do-ministerio-da-saude-expoe-dados-de-16-milhoes-de-pacientes-de-covid/>.

3 PRIVACIDADE DIFERENCIAL: TEORIA E APLICAÇÕES

3.1 REVISÃO LITERÁRIA

Antes da formulação de DP, inclusive do século XX, já existiam áreas de pesquisa voltadas à privacidade. Amplamente difundida, a SDC *(Statistical Disclosure Control*[1]*)* estuda diferentes abordagens para liberar dados sem expor a identidade individual. Ela foi "estatisticamente" iniciada por Fellegi (1972), o qual define privacidade como o direito de determinar qual informação a nosso respeito compartilharemos aos outros. Posteriormente, Dalenius (1977) reconhece e alerta que agências estatísticas precisariam reestruturar seus métodos de proteção para publicar seguramente dados confidenciais. Fato este, formalizado mais tarde (DUNCAN; LAMBERT, 1986).

Conceitos usuais de SDC incluem k-anonimato (SWEENEY, 2002) e sua extensão L-diversidade (MACHANAVAJJHALA *et al.*, 2007) proposta para suprir vulnerabilidades. Basicamente, essas técnicas fornecem proteção por meio da redução e perturbação das informações contidas no banco de dados. A redução abrange soluções triviais mas pouco eficazes como: remoção de variáveis/observações muito sensíveis (supressão), arredondamento de valores para centenas mais próximas (generalização), agrupamento de variáveis por intervalos em que se divulgam categorias genéricas (por exemplo, ao invés de especificar o bairro de um estudo, informe somente a cidade ou região). Já a perturbação de dados fornece garantias de privacidade mais resistentes. A intenção é tornar os indivíduos indistinguíveis inserindo aleatoriedade no processo gerador dos dados, porém, mantendo a representatividade no estudo. Essa adição aleatória pode ser imposta em diferentes etapas do processo de estimação, a qual será discutida de forma mais aprofundada na Seção 4.2.

Há uma vasta literatura no contexto de SDC que alega desassociar indivíduos de um conjunto de dados através da anonimização baseada na substituição de identificadores característicos. Para mais detalhes, ver Adam e Worthmann (1989); Willenborg e Waal (2012). Apesar destas técnicas alcançarem privacidade, sua completa efetividade depende de vários aspectos como o tipo dos dados, suas especificidades (contínuo, discreto, nominal, ordinal) e o nível limitante do risco de divulgação (*disclosure risk*). Adicionalmente, muitos métodos de anonimização mostraram-se vulneráveis a ataques de re-identificação (OHM, 2009; ROCHER; HENDRICKX; MONTJOYE, 2019) e – ao contrário de DP – não compreendem uma rigorosa definição que permita quantificar o grau de privacidade.

[1] Controle Estatístico de Divulgação.

Como exemplo ilustrativo na Tabela 1, apresentamos dados fictícios adaptados de Zigomitros *et al.* (2020) contendo 12 indivíduos e suas informações pessoais. O intuito é mostrar as possibilidades de re-identificação individual em bancos de dados que passaram por métodos tradicionais de anonimização (Tabela 2). Entre eles, o k-anonimato assume que uma tabela é k-anônima se cada registro não puder ser distinguido de pelo menos $k-1$ outros registros de um mesmo atributo. Ou seja, a tabela é particionada em grupos de equivalência com tamanho mínimo k, em que a probabilidade de re-identificar um registro em um grupo é $1/k$. Por conveniência, consideramos as variáveis CEP, nacionalidade e idade sujeitas ao acesso de invasores para ataques de ligação, uma vez que podem ser encontradas publicamente na internet.

A Tabela 2a exibe os dados após serem anonimizados de forma simples com a técnica de supressão em que todos ou alguns valores de uma coluna são omitidos (onde "*" denota um valor suprimido). Mesmo suprimindo a variável de identificação explícita como "Nome", ainda é possível revelar o indivíduo através da combinação das demais variáveis. Por exemplo, com acesso à idade de Raquel, um invasor confirma sua posição na 5ª linha e descobre que foi diagnosticada com câncer. Neste caso, a idade identifica exclusivamente todos pacientes e mesmo suprimindo-a, as combinações de CEP e nacionalidade identificariam a todos, exceto Romeu e Julieta (inevitavelmente expostos ao considerar suas condições de saúde).

Na Tabela 2b, o banco de dados é submetido ao popular k-anonimato com $k = 4$, uma vez que há pelo menos $k-1 = 3$ informações iguais em cada variável. Tal método, geralmente recategoriza os atributos (CEP e idade) e suprime os demais, como é o caso da nacionalidade que não pôde ser categorizada sem violar o 4-anonimato. Muitos ataques adversários priorizam a busca por informações sensíveis antes da identificação individual. Nesse caso, se um invasor conhece a idade e o CEP de Bruce (31 e 13053, respectivamente), descobre que tem câncer pelas categorias 30-40 e 130**, não importando em qual das últimas quatro linhas ele está situado. Além disso, informações auxiliares sob domínio do invasor potencializam a exposição dos dados. Por exemplo, supondo que o CEP 14853 é conhecido, restam 4 pessoas e 3 condições de saúde. Ao saber que o indivíduo-alvo possui uma dieta com restrições de sódio, pode-se inferir que ele sofre de hipertensão. Portanto, enfatiza-se que o método k-anonimato é vulnerável contra ataques envolvendo conhecimentos externos (GANTA; KASIVISWANATHAN; SMITH, 2008).

Alternativamente, a geração de dados sintéticos é uma área de pesquisa usualmente empregada para liberar dados privados a partir de um modelo. Em suma, o conjunto de dados

Tabela 1 – Tabela de dados hipotéticos de saúde

Nome	CEP	Idade	Nacionalidade	Condição
João	13053	28	Russa	Hipertensão
Maria	13068	29	Americana	Hipertensão
Lucas	13068	21	Japonesa	Infecção viral
Ana	13053	23	Americana	Infecção viral
Raquel	14853	50	Indiana	Câncer
Paulo	14853	55	Russa	Hipertensão
Romeu	14850	47	Americana	Infecção viral
Julieta	14850	49	Americana	Infecção viral
Bruce	13053	31	Americana	Câncer
Clark	13053	37	Indiana	Câncer
Tony	13068	36	Japonesa	Câncer
Diana	13068	35	Americana	Câncer

Tabela 2 – Tabela hipótetica sob anonimização

(a) Supressão

Nome	CEP	Idade	Nacionalidade	Condição
*	13053	28	Russa	Hipertensão
*	13068	29	Americana	Hipertensão
*	13068	21	Japonesa	Infecção viral
*	13053	23	Americana	Infecção viral
*	14853	50	Indiana	Câncer
*	14853	55	Russa	Hipertensão
*	14850	47	Americana	Infecção viral
*	14850	49	Americana	Infecção viral
*	13053	31	Americana	Câncer
*	13053	37	Indiana	Câncer
*	13068	36	Japonesa	Câncer
*	13068	35	Americana	Câncer

(b) k-anonimato

Nome	CEP	Idade	Nacionalidade	Condição
*	130**	< 30	*	Hipertensão
*	130**	< 30	*	Hipertensão
*	130**	< 30	*	Infecção viral
*	130**	< 30	*	Infecção viral
*	148**	> 40	*	Câncer
*	148**	> 40	*	Hipertensão
*	148**	> 40	*	Infecção viral
*	148**	> 40	*	Infecção viral
*	130**	30-40	*	Câncer
*	130**	30-40	*	Câncer
*	130**	30-40	*	Câncer
*	130**	30-40	*	Câncer

original (confidencial) é omitido e divulga-se apenas dados novos processados pelo computador de forma "artificial" (RUBIN, 1993). Após a modelagem estatística, gera-se ocorrências aleatórias de uma distribuição de probabilidade com parâmetros adequados que preserve os padrões relevantes dos dados verdadeiros e descreva o comportamento da variável de interesse. Desse modo, a acurácia das análises ou tarefas de classificação é mantida. Em alguns casos, o desenvolvimento de um banco de dados sintéticos baseia-se em um processo de decodificação em que para obter a resposta verdadeira aproximada, aplica-se um tipo de transformação nas observações como a multiplicação por um fator de escala (KINGMA; WELLING, 2013).

De modo complementar, o método de imputação múltipla comumente utilizado para suprir dados faltantes pode ser adotado na criação de dados sintéticos, possibilitando a divulgação destes sem violações de privacidade (RAGHUNATHAN; REITER; RUBIN, 2003). A relevância de DP aliada à geração de dados sintéticos, resultou no popular desafio tecnológico[2] proposto em 2018. Naquela ocasião, um total de 150 mil dólares em prêmios foi concedido a

[2] Mais informações em <https://www.challenge.gov/challenge/differential-privacy-synthetic-data-challenge>.

participantes que criaram matematicamente um gerador de dados sintéticos que atendesse as rigorosas garantias diferencialmente privadas. Algoritmos desse porte flexibilizam a simulação de novos dados com tipos/tamanhos diversificados, conjuntamente ao rigor teórico e fornecem segurança ao compartilhar um conjunto de dados sintéticos.

Para o alcance de privacidade é importante que a inferência seja realizada sobre a população em geral e não individualmente, assim o interesse reside em compreender o comportamento de uma forma coletiva. É preciso avaliar o impacto da restrição de privacidade no desempenho do algoritmo e estudar um *trade-off* (equilíbrio) entre a utilidade dos dados compartilhados com o risco de divulgação. Tecnicamente, tal relação se dará entre acurácia $\times$ perda de privacidade, ponderando o grau de aleatoriedade no modelo para não-identificação dos indivíduos e mantendo a representatividade estatística dos dados. Esta relação é mensurável e viável com a utilização de DP – como enfatiza e compara Jain, Gyanchandani e Khare (2018):

> "Privacidade diferencial, ao contrário de esquemas anteriores, é atualmente a proteção de privacidade mais forte, a qual não requer precauções e suposições sobre informações prévias dos invasores. A comunidade acadêmica prefere privacidade diferencial devido a sua forte limitação matemática ao vazamento de privacidade. A primeira geração de proteção à privacidade aplica técnicas para remoção ou substituição de identificadores sensíveis explícitos dos clientes. A segunda geração realiza a sanitização de bancos de dados com certa anonimização/diversidade e a terceira geração é DP" (JAIN; GYANCHANDANI; KHARE, 2018, p. 20, tradução nossa).

Dentre os estudos que motivaram o desenvolvimento de DP, destaca-se Denning e Denning (1979). Os autores verificaram ao analisar um banco de dados (supostamente privado) que era possível expor informações particulares dos indivíduos com algumas consultas direcionadas. Impactante foi compreender que a privacidade poderia ser preservada se cada nova consulta aos dados não contribuísse à consultas anteriores. Ainda, Dinur e Nissim (2003) demonstra que é necessário um número ainda menor de consultas aleatórias para revelar o conteúdo do banco de dados privado através de um algoritmo de reconstrução. O desfecho do trabalho prova que para alcançar a privacidade, os dados devem ser suficientemente "perturbados", ou seja, adicionar ruído em uma determinada quantidade. No artigo de Dwork *et al.* (2006), definições matemáticas são introduzidas para formalizar a garantia de privacidade conhecida como *differential privacy*. A ideia é fornecer para cada indivíduo aproximadamente a mesma privacidade que resultaria da omissão de seus dados, ou seja, garante indistinguibilidade contra invasores e que nenhum dado individual tenha influência evidente nas informações divulgadas sobre o conjunto de dados (DWORK *et al.*, 2006).

Há um grande repertório de técnicas de aprendizado estatístico disponíveis na literatura a qual DP está inserido. Algumas das linhas de pesquisa que a consideram: regressão regularizada (Ridge e Lasso) (DANDEKAR; BASU; BRESSAN, 2018), análise de componentes principais (JIANG *et al.*, 2013; CHAUDHURI; SARWATE; SINHA, 2013), árvores de decisão (FRIEDMAN; SCHUSTER, 2010), processamento de sinais (SARWATE; CHAUDHURI, 2013), classificadores multi-classes (PATHAK; RAJ, 2012), testes de associação genética (YU *et al.*, 2014), inferência bayesiana (ZHANG; RUBINSTEIN; DIMITRAKAKIS, 2015), sistemas de recomendação (MCSHERRY; MIRONOV, 2009), *deep learning* (ABADI *et al.*, 2016), entre outros.

3.2 APLICAÇÕES PRÁTICAS

Atualmente, observa-se uma forte dedicação no desenvolvimento de novas técnicas para alcançar DP. Empresas de tecnologia como Microsoft, Apple, Google, IBM e SAP têm investido recursos nessa área, bem como universidades que recebem financiamento público ou privado. Para demonstrar a utilidade prática, destacamos algumas abordagens na indústria que incorporam DP em suas práticas:

Google desenvolveu um método chamado RAPPOR (*Randomized Aggregatable Privacy-Preserving Ordinal Response*) que foi implementado em um software *end-user* embutido no navegador Chrome para obtenção de estatísticas dos usuários sob comprometimento de confidencialidade individual (ERLINGSSON; PIHUR; KOROLOVA, 2014).

Apple implementou uma abordagem DP local no iPhone para aperfeiçoar a recomendação de emojis, sugestão de palavras *(QuickType)*, dicas de pesquisa nas notas e no sistema de busca Spotlight. Ainda, no macOS utilizam DP para melhorar a autocorreção de palavras (APPLE, DIFFERENTIAL PRIVACY TEAM, 2017).

Uber lançou um algoritmo em código aberto[3] que impõe DP em consultas via linguagem de consulta estruturada (SQL) para permitir a sua equipe analisar os dados dos usuários, padrões de tráfego e receita dos motoristas. A abordagem baseia-se na sensibilidade elástica que determina a quantidade de ruído exigida para tornar o resultado de uma consulta diferencialmente privado (JOHNSON; NEAR; SONG, 2018).

IBM oferece um conjunto de ferramentas agregadas em uma biblioteca programada na linguagem Python para aprendizado de máquina e análise de dados com garantias de

[3] Disponível em <https://github.com/uber/sql-differential-privacy>.

privacidade incorporadas (HOLOHAN *et al.*, 2019). Ainda, no intuito de aperfeiçoar medidores inteligentes, pesquisadores do IBM desenvolveram um sistema que preserva a privacidade para integrar dispositivos IoT (LYU *et al.*, 2018).

Um marco nesta linha de pequisa, deu-se pela adoção de DP no censo americano de 2020 como metodologia de prevenção contra ataques adversários (ABOWD, 2018). Ao longo das décadas, o Census Bureau dos Estados Unidos opera sob legislação sigilosa das respostas dos questionários. Contudo, em 2020 foi estabelecida a utilização de uma estrutura diferencialmente privada para divulgação completa dos dados sem revelar informações confidenciais dos entrevistados. A decisão foi motivada quando uma pesquisa interna em censos anteriores revelou vulnerabilidades graves expostas pelo método de reconstrução de dados proposto por Dinur e Nissim (2003). Algumas das metodologias diferencialmente privadas sendo implementadas no censo estão documentadas em Li *et al.* (2015a) e Machanavajjhala *et al.* (2008). Outro importante projeto promovido pela universidade de Harvard, o *Privacy Tools Project*[4] objetiva disseminar o entendimento e criação de ferramentas para privacidade de dados. Este é financiado por grandes instituições e estende-se ao longo de várias unidades e universidades.

Um banco de dados privado coletado sob promessa de confidencialidade, mesmo ao passar por processos de anonimização pode ser desprotegido e exposto quando submetido a métodos de reconstrução (DINUR; NISSIM, 2003; ROCHER; HENDRICKX; MONTJOYE, 2019). Alguns trabalhos realizam a re-identificação aplicados à bancos de dados reais: Narayanan e Shmatikov (2008) trabalham no banco dados da Netflix liberado sob simples anonimização para um concurso premiado em US$ 1 milhão ao criador do sistema de recomendação mais eficaz[5]. Através da técnica proposta, os autores demonstram como um invasor poderia facilmente identificar indivíduos com algumas classificações vinculadas de filmes e suas datas aproximadas de visualização. Ainda neste banco de dados, McSherry e Mironov (2009) abordam como fornecer garantias de privacidade sem distorcer excessivamente as inferências. O algoritmo proposto pelos autores é adaptado a uma estrutura DP que neutraliza ataques de ligação da Netflix com os demais perfis dos usuários.

3.3 GARANTIAS

Recentemente DP tornou-se popular nos meios acadêmicos e nas empresas de tecnologia. Grande parte dos pesquisadores que trabalham com privacidade e anonimato voltaram

[4] Para mais detalhes, ver <https://privacytools.seas.harvard.edu>.
[5] Informações adicionais em <https://www.netflixprize.com>.

suas atenções à essa definição atraente de alcançar privacidade através da adição de ruído. Cito os principais motivos:

1. **Proteção das informações contra ataques invasivos**: Ao empregar as definições antecedentes a DP, era preciso antecipar possíveis ataques projetando os objetivos dos invasores para aplicar técnicas de modo a evitá-los (GANTA; KASIVISWANATHAN; SMITH, 2008). DP neutraliza ataques de ligação e garante confidencialidade contra riscos arbitrários, independente de conhecimentos prévios ou informações auxiliares que o invasor possa ter (outros estudos, dados, sites, notícias, etc.).
2. **Quantificação da perda de privacidade**: Não havia uma relação clara com parâmetros para compreender quão privado um conjunto de dados é. Uma das vantagem de DP comparativamente a seus concorrentes é a possibilidade de mensurar a proteção de privacidade. Com embasamento formal é possível desenvolver abordagens consistentes que sustentem a quantia de perda de privacidade com base no ruído adicional. Ainda, é importante no quesito de controle, para monitorar quanta informação está sendo liberada a quem tem acesso aos dados.
3. **Composição de mecanismos**: Permite combinar diversos algoritmos diferencialmente privados agregando-os. A composição é uma propriedade que enfraquece o grau de proteção mas mantém a privacidade. Representa uma maneira de manter o controle do nível de risco conforme amplia-se a complexidade do modelo e mais consultas aos dados são realizadas.

DP é fundamentada na geração de informações sobre a população através de um banco de dados sem realizar inferências pontuais sobre um indivíduo, isto é, sem comprometer sua privacidade. Ela assegura que qualquer resposta a uma determinada consulta tenha ocorrência igualmente provável, independente da presença ou ausência de um indivíduo no banco de dados. Esta justificativa pode ser pertinente em uma pesquisa ao convencer os indivíduos a participarem genuinamente. Portanto, da perspectiva do pesquisador é plausível argumentar: "a abordagem DP garante por definição que as conclusões serão obtidas com quase mesma probabilidade independente de suas respostas, ou seja, o que descobriremos com você, será muito similar sem você."

É importante esclarecer que DP não é um algoritmo em si, é uma definição que deve ser interpretada como um conjunto de condições formais a serem satisfeitas e avaliadas para conferir as propriedades de privacidade de um algoritmo. A aleatoriedade contida nos algoritmos diferencialmente privados é essencial para assegurar privacidade. Se a função for determinística e

não aleatória, as garantias de privacidade não serão satisfeitas. Contudo, a introdução de incerteza deve ser realizada cautelosamente pois grandes quantidades de aleatoriedade aumentarão a proteção, mas comprometerão a utilidade do modelo. Idealmente, a adição de ruído deve ser suficiente para impedir ataques adversários de modo que não viese demasiadamente os estimadores.

3.4 DEFINIÇÃO FORMAL

Antes de formalizar DP, é necessário definir algumas medidas e notações preliminares. Seja um conjunto de dados $D = \{x_1, x_2, \ldots, x_n\} \in \mathscr{D}^n$ contendo n observações. A distância entre dois conjuntos de dados D e D' é calculada pela norma-ℓ_1, i.e., $||D - D'||$ em que $||D||_1 = \sum_{i=1}^{n} |x_i|$. Quando os dados possuem mesmo tamanho e são dispostos por múltiplas linhas, esta diferença é chamada distância de Hamming. Denotaremos $D \sim D'$ por bancos de dados adjacentes se D e D' diferem na contribuição de um único elemento, ou seja, se a distância de Hamming for igual a 1. Por exemplo, se $D = \{1,2,3,4,5\}$ e $D' = \{2,3,4,5,6\}$ temos que $D \sim D'$. Um algoritmo diferencialmente privado é fundamentado na proteção da privacidade entre bancos de dados adjacentes diferindo em no máximo um indivíduo[6]. Dessa forma, ao analisar a saída do algoritmo não é possível identificar a inclusão ou exclusão de um indivíduo em particular (GANTA; KASIVISWANATHAN; SMITH, 2008).

Um mecanismo de privacidade é uma função aleatória que mapeia o banco de dados para uma distribuição de probabilidade em algum intervalo e retorna um vetor de números reais escolhidos aleatoriamente dentro do intervalo. A função é diferencialmente privada se a probabilidade de qualquer saída é afetada por um pequeno fator multiplicativo ao adicionar ou remover uma única observação no banco de dados. Portanto, um invasor será incapaz de inferir sobre a participação ou não dos indivíduos. Dados os formalismos, podemos definir DP (DWORK *et al.*, 2006):

Definição 3.1. Uma função aleatória $\mathcal{M} : \mathscr{D} \to S$ fornece ε-DP para todos conjuntos de dados D e D' adjacentes se

$$\mathbb{P}\big[\mathcal{M}(D) \in S\big] \;\leq\; \exp(\varepsilon) \times \mathbb{P}\big[\mathcal{M}(D') \in S\big], \tag{3.1}$$

ou equivalentemente,

$$\left|\log\left(\frac{\mathbb{P}[\mathcal{M}(D) \in S]}{\mathbb{P}[\mathcal{M}(D') \in S]}\right)\right| \;\leq\; \varepsilon, \tag{3.2}$$

[6] Encontra-se abordagens variantes na literatura que consideram diferentes tamanhos para os bancos de dados (HSU *et al.*, 2014).

em que $\mathbb{P}(\cdot)$ é uma medida de probabilidade, ε é um número real positivo e S são todos os subconjuntos de possíveis resultados no domínio de $\mathcal{M}$. A probabilidade de obter S no banco de dados D é quase a mesma quando os dados são D'. Tal probabilidade é alterada por no máximo um fator multiplicativo $\exp(\varepsilon)$. Denomina-se ε como uma medida de orçamento de privacidade (*privacy budget*), mais conhecida como perda de privacidade (*privacy loss*). Assim, menores valores de ε correspondem a mais privacidade, uma vez que a chance de identificação é reduzida. Note de (3.2) que

$$-\varepsilon \;\leq\; \log\left(\frac{\mathbb{P}\left[\mathcal{M}(D)\in S\right]}{\mathbb{P}\left[\mathcal{M}(D')\in S\right]}\right) \;\leq\; \varepsilon,$$

ou seja, o logaritmo das chances de obter uma resposta específica ao incluir ou remover uma única observação está entre $[-\varepsilon,\ \varepsilon]$.

A literatura apresenta disparidades na escolha apropriada de ε para uso prático (LEE; CLIFTON, 2011). Há alguns fatores que influenciam como a disposição dos dados e grau de sigilo desejado. Usualmente, o valor do parâmetro ε varia de 0.01 a 10. Dwork (2008) declara que a seleção dessa constante depende do contexto, porém, sugere valores entre 0.01 e até 3 em alguns casos. Basicamente, um valor ideal de ε remete a escolher um *trade-off* entre privacidade e acurácia. Recomenda-se adotar uma abordagem rigorosa com $\varepsilon < 1$ pois ao assumir perda de privacidade grande, há maiores riscos de exposição dos dados com práticas intrusivas. De modo complementar, a literatura explora interpretações da definição de DP de uma perspectiva Bayesiana (ABOWD; VILHUBER, 2008; ZHANG; RUBINSTEIN; DIMITRAKAKIS, 2015).

Diferentes formas de "flexibilização" foram desenvolvidas no âmbito de DP. Em geral, estas fornecem garantias de privacidade mais fracas em troca da menor quantidade de ruído imposto no processo. A principal generalização foi proposta ao incrementar um parâmetro de locação δ em (3.1) que fornece um "relaxamento" da restrição de privacidade (DWORK *et al.*, 2006). A extensão (ε,δ)-DP possui uma definição estritamente mais fraca que a original, porém proporciona uma melhor acurácia devido a menor magnitude de ruído adicional (DWORK; SMITH, 2010). Sua origem foi motivada ao verificar que a privacidade era raramente violada tratando-se de eventos distintos com chances de ocorrência improvável. Então, a probabilidade de qualquer indivíduo sofrer perda de privacidade superior a ε foi limitada pelo parâmetro arbitrário δ (mesmo que minimamente). Assim, garante-se que o valor absoluto da perda de privacidade do algoritmo não exceda ε com probabilidade $1-\delta$ (DWORK; ROTH, 2014, Lema 3.17). Quanto mais próximo δ for de zero, mais privacidade será assegurada. Segundo Dwork e Roth (2014), é desejável que δ assuma valores menores que $1/n$. Caso contrário, há o risco de "apenas

algumas[7]" observações serem expostas.

As duas principais definições se distinguem de tal forma: (i) ε-DP garante que para cada execução do mecanismo $\mathcal{M}$, a saída é "quase" igualmente provável de ser observada nos bancos de dados adjacentes, simultaneamente. (ii) (ε, δ)-DP diz que para cada banco de dados adjacentes é extremamente improvável distinguir se a saída gerada via $\mathcal{M}$ foi obtida utilizando o banco de dados D ou D' (DWORK; ROTH, 2014).

Vale apontar outras propostas derivadas da DP pura: A DP local concentra sua metodologia em cada usuário proteger suas informações com mecanismos anonimizadores antes de submetê-los ao coletor de dados (DUCHI; JORDAN; WAINWRIGHT, 2013). Esta abordagem evita problemas com o curador/agregador central dos dados que eventualmente pode ser hackeado ou é não-confiável. A desvantagem de sua aplicação é a baixa utilidade, uma vez que cada usuário adiciona ruído aos seus dados totalizando uma grande contaminação contida no procedimento (CORMODE *et al.*, 2018). Uma das mais recentes extensões desenvolvida na literatura chama-se DP concentrada. A proposta é voltada para mecanismos em que a perda de privacidade é pequena e fortemente concentrada em torno de uma média subgaussiana. Para mais detalhes e definição de uma variável aleatória com distribuição subgaussiana, ver Dwork e Rothblum (2016).

A crítica a versões alternativas de DP argumenta que em se tratando de privacidade qualquer "relaxamento" pode ser catastrófico. Alguns autores afirmam – vide Frank McSherry[8], coautor da DP clássica – que um parâmetro flexível pode causar falhas comprometedoras mesmo definindo-o como um valor ínfimo e salienta a importância de investigar a tolerância de riscos "aparentemente" insignificantes. Ele ainda alega que a escolha de δ (muitas vezes determinada como $1/n$) é uma dependência do número de observações que pode prejudicar a garantia de privacidade. No trabalho de Ganta, Kasiviswanathan e Smith (2008), são mostradas algumas definições mais fracas de DP que falham completamente em proteger dados confidenciais, assim como versões maleáveis que fornecem garantias de privacidade semelhantes à definição tradicional.

---

[7] O conceito *"just a few"* apresenta lacunas tratando-se de privacidade pois "apenas algumas" informações estão desprotegidas. Com essa abordagem, a proteção de um conjunto de dados é priorizada aos indivíduos comuns (maioria) e corre-se um maior risco de exposição para indivíduos com comportamento discrepante (*outliers*). Em muitos casos, faz-se necessário ponderar informações que demandam maior sigilo.

[8] Acesso em <https://github.com/frankmcsherry/blog>.

3.5 PROPRIEDADES

3.5.1 Composição

Os algoritmos diferencialmente privados podem ter múltiplas saídas. A cada resultado divulgado, as informações dos dados "vazam" e, portanto, precisam de proteção adicional já que consultas consecutivas degradam a privacidade (GANTA; KASIVISWANATHAN; SMITH, 2008). Os teoremas de composição formalizam essa ideia considerando a repartição do valor de ε usado para cada saída. É uma propriedade poderosa, pois permite decompor mecanismos DP e combiná-los para obter algoritmos mais sofisticados. A junção de mecanismos DP ou a distribuição conjunta do produto destes também satisfaz DP.

Teorema 3.1. Suponha que um método inclua m mecanismos independentes $\mathcal{M}_1, \ldots, \mathcal{M}_m$ cujas garantias de DP são $\varepsilon_1, \ldots, \varepsilon_m$; ao combiná-los obtém-se um algoritmo $(\sum_{j=1}^{m} \varepsilon_j)$-DP.

Prova: Sejam dois algoritmos independentes $\mathcal{M}_1 : \mathbb{N} \to \mathcal{R}_1$ e $\mathcal{M}_2 : \mathbb{N} \to \mathcal{R}_2$ com restrições de privacidade ε_1-DP e ε_2-DP, respectivamente. A combinação destes satisfaz $(\varepsilon_1 + \varepsilon_2)$-DP. Considere qualquer par de pontos $(r_1, r_2) \in \mathcal{R}_1 \times \mathcal{R}_2$, então:

$$\begin{aligned}
\frac{\mathbb{P}[\mathcal{M}_{1,2}(D) = (r_1, r_2)]}{\mathbb{P}[\mathcal{M}_{1,2}(D') = (r_1, r_2)]} &= \frac{\mathbb{P}[\mathcal{M}_1(D) = r_1]\,\mathbb{P}[\mathcal{M}_2(D) = r_2]}{\mathbb{P}[\mathcal{M}_1(D') = r_1]\,\mathbb{P}[\mathcal{M}_2(D') = r_2]} \\
&= \left(\frac{\mathbb{P}[\mathcal{M}_1(D) = r_1]}{\mathbb{P}[\mathcal{M}_1(D') = r_1]}\right)\left(\frac{\mathbb{P}[\mathcal{M}_2(D) = r_2]}{\mathbb{P}[\mathcal{M}_2(D') = r_2]}\right) \\
&\leq \exp(\varepsilon_1)\exp(\varepsilon_2) \\
&= \exp(\varepsilon_1 + \varepsilon_2).
\end{aligned}$$

Por indução, é válido para $\mathcal{M}_1, \ldots, \mathcal{M}_m$ garantindo $(\sum_{j=1}^{m} \varepsilon_j)$-DP.

Denomina-se ataque de composição quando um invasor busca a exposição individual através da combinação de informações publicadas independentemente. Por exemplo, se um enfermo trata-se em hospitais diferentes e estes divulgam tabelas anonimizadas protegendo individualmente os pacientes. Ainda assim, é possível descobrir sua identidade relacionando ambas informações fornecidas pelos hospitais. A propriedade de composição contribui para impedir a degradação da privacidade resultante de referências cruzadas e permitir liberações privadas independentes (GANTA; KASIVISWANATHAN; SMITH, 2008). É importante mencionar que há segmentações de composição que fogem do escopo deste trabalho. Para mais informações, ver Dwork e Lei (2009).

3.5.2 Pós-processamento

Teorema 3.2. Seja uma função aleatória arbitrária $f : S \to \mathcal{R}$ definida sobre a imagem do mecanismo $\mathcal{M} : \mathscr{D} \to S$ que satisfaz ε-DP, então $f(\mathcal{M}) : \mathscr{D} \to \mathcal{R}$ é ε-diferencialmente privado (DWORK; ROTH, 2014, p. 19).

Essa propriedade permite aplicar transformações independentes a um resultado originado de um mecanismo diferencialmente privado sem afetar suas garantias de privacidade. Quando os dados estão protegidos sob definição de DP, é impossível alterar o algoritmo (com uma função) para torná-lo menos diferencialmente privado. Tal circunstância, é conhecida como robustez/imunidade ao pós-processamento. Dwork e Roth (2014) complementa com uma analogia: "Um analista de dados é incapaz de aumentar a perda de privacidade de um banco de dados sob formalismo de DP, não importa quais informações auxiliares lhe estejam disponíveis" (DWORK; ROTH, 2014, p. 233, tradução nossa).

O pós-processamento auxilia na construção de mecanismos complexos e sua utilização é frequente na área de dados sintéticos para obtenção de dados representativos que não sejam os verdadeiros. A geração destes pode garantir privacidade ao definir os parâmetros de uma distribuição baseados em uma abordagem diferencialmente privada. Assim, pelo Teorema 3.2. de pós-processamento, quaisquer geração de dados utilizando os parâmetros perturbados (obtidos de um modelo DP) serão diferencialmente privados (BOWEN; LIU, 2020). Ainda, o censo americano beneficia-se da aplicação do pós-processamento para a liberação segura de dados (ABOWD, 2018).

3.6 MECANISMOS

No geral, define-se mecanismo como um algoritmo que toma como entrada um banco de dados e retorna uma resposta que pode ser um número, uma estatística, um coeficiente, entre outros. Os mecanismos diferencialmente privados são responsáveis por preservar a privacidade dos indivíduos através da adição de ruído aleatório para mascarar os dados e fornecer confidencialidade. Deve-se estar ciente que a garantia de privacidade reduzirá a acurácia, uma vez que incerteza é incorporada no modelo. Contudo, é preciso ter cautela ao contaminar os dados. Primeiro, pois se a perturbação não for adequada e suficiente para alcançar o anonimato, as informações estarão vulneráveis à exposição por meio de ataques invasivos. Segundo, porque ruído em quantidade desmedida irá gerar resultados altamente viesados e arruinará a utilidade/representatividade do modelo.

3.6.1 Sensibilidade

Em geral, um mecanismo que satisfaz DP calibra a quantidade de ruído com base na mudança máxima possível entre bancos de dados adjacentes $D \sim D'$ para descobrir a influência causada pela presença ou não de um indivíduo. Essa mudança, denominada sensibilidade, é a principal métrica para mensurar a quantia de aleatoriedade necessária no algoritmo para alcançar privacidade omitindo a participação exclusiva de cada indivíduo.

Definição 3.3. Para qualquer função $f : \mathscr{D} \to \mathbb{R}^d$, a sensibilidade de f para conjuntos de dados adjacentes D, D' é dada por

$$\Delta f = \max_{D \sim D'} ||f(D) - f(D')||_1,$$

em que $\mathscr{D}$ é o conjunto de todos bancos de dados possíveis (DWORK *et al.*, 2006). Em particular, quando $d = 1$, a sensibilidade mede a diferença entre os valores que a função f pode assumir entre qualquer par de bancos de dados que diferem em apenas um elemento. Em termos estatísticos, Δf captura o efeito que uma única observação do banco de dados pode alterar no resultado calculado pela norma-ℓ_1. Quanto maior for o valor de Δf, maior será o risco de re-identificação (*disclosure risk*) dos indivíduos. Dessa forma, sensibilidade alta requer acentuada adição de ruído para proteção, enquanto a baixa sensibilidade demanda menos perturbação nos dados verdadeiros.

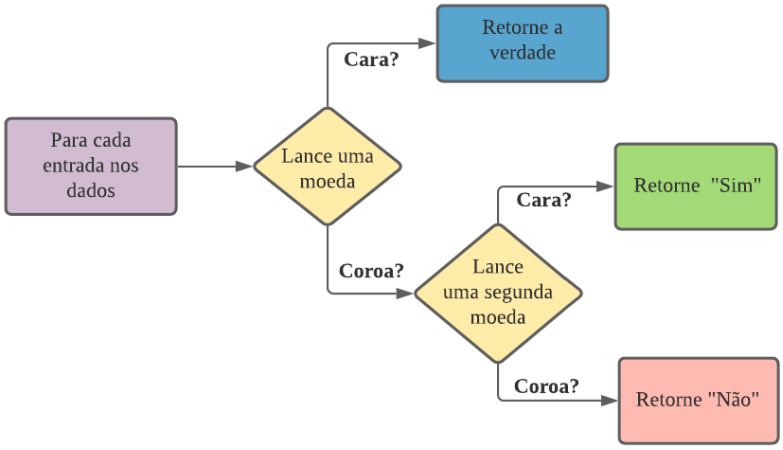
Há variantes para mensurar a sensibilidade, sendo esta uma escolha que reflete no desempenho dos métodos diferencialmente privados (AWAN; SLAVKOVIĆ, 2020). Por exemplo, em aplicações de minimização do risco empírico é usual adotar a norma-ℓ_2 (Euclidiana) (CHAUDHURI; MONTELEONI; SARWATE, 2011; KIFER; SMITH; THAKURTA, 2012). Ainda, o método *smooth sensitivity* tenta aproximar funções com alta sensibilidade apenas no "pior caso" adicionando ruído suavizado com magnitude determinada pela função e também pelo banco de dados. Embora esse método exija uma quantidade menor de ruído para alcançar privacidade, a sensibilidade *smoothed* é de difícil obtenção analítica (NISSIM; RASKHODNIKOVA; SMITH, 2007).

Após estabelecer a magnitude do ruído necessário para garantir DP por meio da sensibilidade e do parâmetro ε, é preciso determinar mecanismos adequados que forneçam confidencialidade a identidade dos indivíduos. Como já mencionado, a privacidade é alcançada com a presença de incerteza, a qual é obtida com aleatoriedade inserida nos procedimentos analíticos que constituem mecanismos diferencialmente privados.

3.6.2 Mecanismo de resposta aleatória

Motivada por cientistas sociais, a técnica *randomized response* $\mathcal{M}_R$ foi desenvolvida na década de sessenta para coletar dados sobre comportamentos ilegais ou constrangedores que sofriam resistência diante dos entrevistados (WARNER, 1965). Naquela época, foi o recurso adotado pelos pesquisadores como garantia de sigilo aos participantes. O fluxograma adaptado (Figura 2) mostra como a aleatoriedade é inserida no mecanismo para mascarar as respostas dos indivíduos:

Figura 2 – Fluxograma do mecanismo de resposta aleatória.



O diagrama é um exemplo para compreender como a privacidade é obtida por um mecanismo de escolha aleatória, ou seja, a interpretação das respostas é influenciada pelo processo a qual foram submetidas. Neste contexto, há um termo adequado "*plausible deniability*" indicando que qualquer saída liberada pelo mecanismo pode ser "plausivelmente negada". Assim, um invasor é impedido de inferir se um registro específico do banco de dados foi mais/menos responsável por gerar uma saída observada.

A principal vantagem obtida com o mecanismo caracteriza-se pela coleta das respostas já perturbadas, ou seja, as verdadeiras informações pessoais estão mascaradas mesmo sob ataques. Logo, o emprego de $\mathcal{M}_R$ é ideal para usuários que desejam ocultar seus dados do próprio coletor de dados. A desvantagem é que o mecanismo tem seu desempenho dependente do tamanho amostral (melhora conforme o aumento do número de participantes) (LENSVELT-MULDERS; HOX; HEIJDEN, 2005). O RAPPOR (citado na Seção 3.2) representa um exemplo popular de aplicação em privacidade baseado no mecanismo de resposta aleatória (ERLINGSSON; PIHUR; KOROLOVA, 2014).

3.6.3 Mecanismo de Laplace

No âmbito de DP, a privacidade é usualmente obtida via mecanismo de Laplace. Denotado por $\mathcal{M}_L$, o mecanismo consiste em introduzir incerteza sobre os elementos ou resultado de uma análise. Essa perturbação é realizada pela adição de ruído aleatório gerado da distribuição de Laplace a fim de mascarar os dados verdadeiros (DWORK *et al.*, 2006). Seja X uma variável aleatória com distribuição de Laplace de parâmetros μ e b (ruído de Laplace). Esta distribuição foi proposta por Pierre-Simon Laplace (1774) como uma versão simétrica da distribuição exponencial, conhecida por exponencial dupla, uma vez que, compõe duas distribuições com suporte positivo e negativo. Sua função densidade de probabilidade é dada por

$$\text{Lap}(x;\mu,b) = \frac{1}{2b}\exp\left(-\frac{|x-\mu|}{b}\right), \tag{3.3}$$

em que μ é o parâmetro de locação e b refere-se ao parâmetro de escala. A média e variância de X são, respectivamente: $\mathbb{E}(X) = \mu$ e $\text{var}(X) = 2b^2$.

A depender dos valores dos parâmetros, a densidade de Laplace assume diferentes formas. A Figura 3 apresenta a função densidade de probabilidade de Laplace centrada em zero ($\mu = 0$), variando o parâmetro de escala b. Observa-se que b controla o encolhimento da distribuição do ruído. O mecanismo satisfaz ε-DP ao adicionar o ruído de Laplace com parâmetro $b = \Delta f/\varepsilon$. Ao considerar a sensibilidade Δf igual a 1, obtém-se como orçamento de privacidade para $b = \{1, 2, 4\}$, os valores de $\varepsilon = \{1, 0.5, 0.25\}$, respectivamente. Note que para menores ε, o grau de privacidade é mais forte, visto que o suporte da distribuição é ampliado/achatado e há mais variabilidade. Em contrapartida, para ε maiores, a densidade da variável aleatória é altamente concentrada em torno de zero e portanto a perturbação deste ruído será menos influente (menos privada).

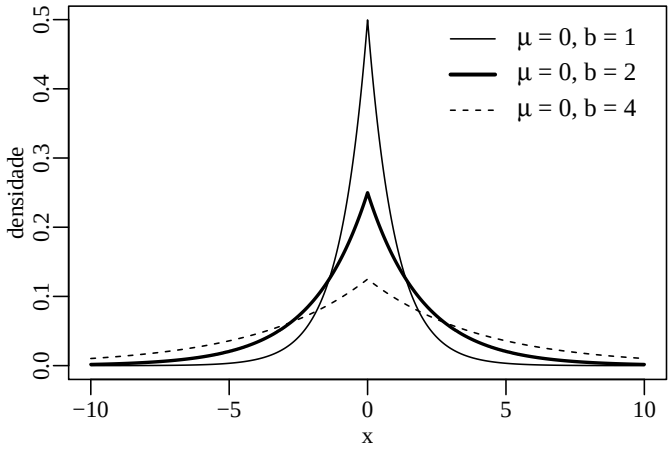
Em mecanismos diferencialmente privados, usualmente adota-se o parâmetro $\mu = 0$ em (3.3) para adição do ruído, a qual será denotada por Lap(b). Para qualquer função $f : \mathscr{D} \to \mathbb{R}$, o mecanismo de Laplace é expresso por

$$\mathcal{M}_L(D) = f(D) + \boldsymbol{z},$$

em que $\boldsymbol{z}$ é um vetor de ocorrências de variáveis aleatórias i.i.d. seguindo uma lei de Laplace $\text{Lap}(\Delta f/\varepsilon)$. A dimensão de $\boldsymbol{z}$ varia conforme o tipo de perturbação empregada, a qual será detalhada na Seção 4.2.

Pela distribuição de Laplace, sabe-se que valores mais próximos dos dados verdadeiros têm maiores probabilidades de serem gerados do que os mais distantes, pois as caudas

Figura 3 – Função densidade de probabilidade de Laplace.



são exponenciais em ambas direções e a medida que nos afastamos do centro da distribuição, as suas probabilidades decrescem a uma taxa exponencialmente rápida $\exp(\varepsilon)$. Indica-se pela variância da distribuição que com sensibilidade baixa e perda de privacidade alta, obteremos valores próximos de zero e a pertubação será de pouca influência. Já, para Δf grande e ε pequeno, o ruído estará mais distante do resultado verdadeiro com maior probabilidade. Assim, a distribuição corresponde aos requisitos de um mecanismo diferencialmente privado e mostra-se apropriada para aplicações em DP.

Teorema 3.3. O mecanismo que adiciona ruído de forma independente, gerado da distribuição de Laplace com parâmetros $\mu = 0, b = \Delta f/\varepsilon,$ satisfaz ε-DP.

Prova: Sejam D e D' conjuntos de dados adjacentes e uma função $f : \mathbb{N} \to \mathbb{R}^n$. A função densidade de probabilidade de $\mathcal{M}_L$ é dada por $\mathbb{P}[\mathcal{M}(D) = r] = \exp(-\varepsilon|f(D) - r|/\Delta f)$ e denotada por $p_D(r)$. De forma análoga para quando o banco de dados é D'. Seja $r_i \in \mathbb{R}$ igual ao i-ésimo ponto r com ruído gerado independentemente, temos que

$$\begin{aligned}
\frac{p_D(r)}{p_{D'}(r)} &= \frac{\prod_{i=1}^{n} \exp\left(-\frac{\varepsilon|f(D)_i - r_i|}{\Delta f}\right)}{\prod_{i=1}^{n} \exp\left(-\frac{\varepsilon|f(D')_i - r_i|}{\Delta f}\right)} \\
&\leq \prod_{i=1}^{n} \exp\left(\frac{\varepsilon|f(D)_i - f(D')_i|}{\Delta f}\right) \\
&= \prod_{i=1}^{n} \exp\left[\frac{\varepsilon(|f(D')_i - r_i| - |f(D)_i - r_i|)}{\Delta f}\right] \\
&= \exp\left(\frac{\varepsilon\,||f(D) - f(D')||_1}{\Delta f}\right) \\
&\leq \exp(\varepsilon),
\end{aligned}$$

em que a primeira expressão é valida ao aplicar a desigualdade triangular ($|a| - |b| \leq |a - b|$) no

expoente. Pela definição de sensibilidade, $||f(D) - f(D')||_1 \leq \Delta f$ e portanto, a razão é limitada por $\exp(\varepsilon)$, satisfazendo o requisito ε-DP.

Como mecanismo pioneiro, o ruído laplaciano tem sido amplamente utilizado para preservar a privacidade (SARATHY; MURALIDHAR, 2011; ZHANG *et al.*, 2012; PHAN *et al.*, 2017). Destaca-se o algoritmo PMW - *Private Multiplicative Weights* (pesos multiplicativos privados), o qual pondera iterativamente os dados perturbados baseado em sua contribuição no resultado (ganhos em acurácia, significância, entre outros). Ou seja, consultas nos dados verdadeiros cuja resposta está muito longe quando o mesmo é feito nos dados privados são ponderadas negativamente. Caso contrário, pesos maiores são atribuídos às observações de tal forma que se obtenha resultados mais acurados (HARDT; ROTHBLUM, 2010).

3.6.4 Outros Mecanismos

Apesar de não utilizá-lo neste trabalho, apresentamos brevemente o mecanismo exponencial ($\mathcal{M}_E$) para conhecimento. Este foi desenvolvido para situações em que se procura a melhor resposta, uma vez que "melhor" se caracteriza pela saída de uma função com máxima qualidade possível (MCSHERRY; TALWAR, 2007). Essa função no contexto da preservação de privacidade é interpretada como a utilidade de um resultado que relaciona um conjunto de alternativas aos dados confidenciais. Portanto, a intenção do usuário é que um mecanismo diferencialmente privado – mesmo perturbando as observações para proteger a privacidade – retorne algum elemento do conjunto de resultados possíveis que maximizem sua utilidade (DWORK; ROTH, 2014).

Em geral, $\mathcal{M}_E$ oferece fortes garantias de utilidade pois à medida que o nível de qualidade diminui, os resultados são exponencialmente ponderados para reduzir sua influência (BOWEN; LIU, 2020). Cito alguns trabalhos influentes que empregam este mecanismo: Hardt, Ligett e Mcsherry (2012) aplica o $\mathcal{M}_E$ em algoritmos aprimorados para identificar respostas mais informativas. Ainda, Wasserman e Zhou (2010) e Snoke *et al.* (2018) trabalham com a geração de dados sintéticos – uma área em que o mecanismo exponencial atua fortemente.

Outros mecanismos foram propostos. Para informações adicionais, ver: mecanismo mediano (ROTH; ROUGHGARDEN, 2010), mecanismo geométrico (GHOSH; ROUGHGARDEN; SUNDARARAJAN, 2012) e mecanismo *staircase* (GENG; VISWANATH, 2015). Destaca-se o mecanismo Gaussiano (DWORK; ROTH, 2014), análogo ao mecanismo de Laplace, porém utiliza a norma-ℓ_2 para medir a sensibilidade e ponderar o ruído (LIU, 2019).

4 MODELOS DE REGRESSÃO

O presente capítulo apresenta inicialmente os modelos tradicionais de regressão linear e logístico. Em seguida, diferentes metodologias de perturbação são definidas para preservação de privacidade DP em bancos de dados. Dado o embasamento, apresentamos os modelos propostos nas versões linear e logística comparando-os com modelos existentes da literatura DP.

4.1 MODELOS USUAIS

Quando a média de uma distribuição estatística não é constante, ou seja, quando o comportamento da variável de interesse é afetado por outras variáveis, comumente utiliza-se modelos de regressão para entender sua influência. Em particular, a regressão linear é amplamente empregada sob condição de erros normalmente distribuídos. Caso não for razoável assumir distribuição normal, como é o caso de variáveis binárias, é possível ajustar uma regressão logística pertencente a classe flexível de modelos lineares generalizados (MLG).

4.1.1 Regressão Linear

Uma das primeiras formas de análise de regressão a ser detalhadamente estudada na literatura foi a linear. Amplamente difundido em aplicações práticas, o modelo linear é assim denominado pois considera que a relação da resposta com os regressores é expressa por uma função linear de parâmetros desconhecidos. A modelagem do comportamento da variável de interesse considerando o efeito de variáveis explicativas é estruturada por

$$\boldsymbol{y} = X\boldsymbol{\beta} + \boldsymbol{u},$$

em que $\boldsymbol{y}$ é um vetor $n \times 1$ das respostas, X é uma matriz de dimensão $n \times k$ $(k < n)$ sendo k a quantidade de variáveis regressoras (covariáveis), $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_k)^\top$ é um vetor k-dimensional de parâmetros desconhecidos, $\boldsymbol{u}$ é um vetor de tamanho $n \times 1$ de erros aleatórios com matriz de covariância $\Omega = \text{diag}\{\sigma_1^2, \ldots, \sigma_n^2\}$. Para viabilizar o processo inferencial, assume-se erros com média zero e não correlacionados. Quando os erros são homoscedásticos (variância constante) temos que $\sigma_i^2 = \sigma^2 > 0 \; \forall_i$, ou seja, $\Omega = \sigma^2 I_n$ em que I_n é a matriz identidade.

A análise de regressão tem como objetivo central fazer inferência sobre $\boldsymbol{\beta}$, uma vez que este vetor representa a influência dos regressores sobre a média da variável resposta, ou seja, $\mathbb{E}(\boldsymbol{y}) = X\boldsymbol{\beta}$. O ajuste do modelo é usualmente realizado via método de mínimos quadrados

ordinários (MQO), o qual consiste em encontrar os valores dos parâmetros que minimizam a soma de quadrados dos erros, dada por

$$\boldsymbol{u}^\top \boldsymbol{u} = (\boldsymbol{y} - X\boldsymbol{\beta})^\top (\boldsymbol{y} - X\boldsymbol{\beta}). \tag{4.1}$$

O estimador MQO de $\boldsymbol{\beta}$ é

$$\widehat{\boldsymbol{\beta}} = \arg\min_{\beta} \left(\boldsymbol{u}^\top \boldsymbol{u} \right) = (X^\top X)^{-1} X^\top \boldsymbol{y}.$$

Esses estimadores possuem propriedades desejáveis como não-viés assumindo $\mathbb{E}(u_i) = 0$, eficiência e consistência, pois convergem em probabilidade para seu verdadeiro valor ($\widehat{\boldsymbol{\beta}} \xrightarrow{p} \boldsymbol{\beta}$). A matriz de covariância do MQO é expressa por

$$\Psi = \operatorname{cov}(\widehat{\boldsymbol{\beta}}) = (X^\top X)^{-1} X^\top \Omega X (X^\top X)^{-1}. \tag{4.2}$$

Na presença de homoscedasticidade, todos os elementos diagonais da matriz Ω são iguais, obtendo-se

$$\Psi = \sigma^2 (X^\top X)^{-1}. \tag{4.3}$$

Dessa forma, a matriz de covariância de $\widehat{\boldsymbol{\beta}}$ pode ser estimada por $\widehat{\sigma}^2 (X^\top X)^{-1}$.

Quando utilizamos um modelo de regressão linear é comum assumir homoscedasticidade, ou seja, variância fixa para todos os erros (ou respostas, consequentemente). Essa pressuposição é restritiva em diversas situações práticas, particularmente na modelagem de dados de corte transversal. Mesmo violada, o estimador usual dos parâmetros MQO permanece não-viesado e consistente, mas deixa de ser o melhor estimador de variância mínima (*BLUE - Best Linear Unbiased Estimator*). Contudo, o estimador tradicional da matriz de covariância (4.3) passa a ser viesado e inconsistente sob heteroscedasticidade. Logo, inferências como estimação intervalar e testes de hipóteses tornam-se pouco confiáveis. Alternativamente, a literatura estuda meios para manter a consistência no processo inferencial e fornecer estimativas válidas em ambos cenários de variabilidade (MACKINNON; WHITE, 1985; NEWEY; WEST, 1987; NEWEY; WEST, 1994; ANDREWS, 1991).

Portanto, além do MQO, consideramos estimadores adequados para estimação consistente da matriz de covariância: o precursor HC0 (WHITE, 1980) e HC4 (CRIBARI-NETO, 2004) sendo uma alternativa mais confiável para construção de testes quasi-*t* em amostras finitas. Em geral, estimadores HCs apresentam bom desempenho sob heteroscedasticidade (LONG; ERVIN, 2000). Essa classe de estimadores consistentes é construída substituindo os elementos diagonais da matriz de covariância Ω na Equação (4.2).

O estimador consistente mais empregado foi proposto por White (1980) e é obtido substituindo o i-ésimo elemento diagonal de $\Omega = \text{diag}\{\sigma_1^2, \ldots, \sigma_n^2\}$ pelo i-ésimo resíduo ao quadrado $\widehat{\Omega}_0 = \text{diag}\{\widehat{u}_1^2, \ldots, \widehat{u}_n^2\}$, dado por

$$\text{HC0} = (X^\top X)^{-1} X^\top \widehat{\Omega}_0 X (X^\top X)^{-1}.$$

Entretanto, foi evidenciado através de resultados numéricos que o HC0 apresenta viés substancial em amostras de tamanho finito e que há distorção inferencial na presença de pontos influentes de alta alavancagem (MACKINNON; WHITE, 1985; CRIBARI-NETO; ZARKOS, 2001). Generalizando o HC0, os estimadores HC1, HC2 (MACKINNON; WHITE, 1985) e HC3 (DAVIDSON; MACKINNON, 1993) foram propostos incorporando termos de correção para melhorar o desempenho na estimação consistente da matriz de covariância de $\widehat{\boldsymbol{\beta}}$ em amostras pequenas. Entre estes, Long e Ervin (2000) conduz um estudo comparativo e constata a superioridade do HC3.

O estimador denominado HC4 (CRIBARI-NETO, 2004) pondera o impacto de pontos com alta alavancagem em amostras finitas no estimador da matriz de covariância e apresenta bom desempenho na construção de testes quasi-t. A correção inclusa considera o grau de alavancagem de cada observação medido pelos elementos diagonais (h_i) da "matriz-chapéu" $H = X(X^\top X)^{-1}X^\top$ e acentua o termo de ajuste baseado no estimador HC3. Define-se HC4 da seguinte forma

$$\text{HC4} = (X^\top X)^{-1} X^\top \widehat{\Omega}_4 X (X^\top X)^{-1},$$

em que

$$\widehat{\Omega}_4 = \text{diag}\left\{ \frac{\widehat{u}_1^2}{(1-h_1)^{\varsigma_1}}, \ldots, \frac{\widehat{u}_n^2}{(1-h_n)^{\varsigma_n}} \right\}, \quad \varsigma_i = \min\left\{ 4, \frac{n h_i}{\mathfrak{a}} \right\},$$

sendo $\mathfrak{a}$ igual a soma de todos níveis individuais de alavancagem, que por sua vez, é igual ao posto da matriz X dado que H é simétrica e idempotente.

Quando há heteroscedasticidade no modelo, consideramos a função cedástica estruturada por

$$\sigma_i^2 = \exp\left\{ \sum_{j=1}^{k} \gamma_j w_{ij} \right\}, \tag{4.4}$$

em que $\boldsymbol{\gamma}$ é um vetor k-dimensional dos parâmetros de dispersão e w_{ij} a covariável associada. Podemos mensurar o grau de heteroscedasticidade por $\lambda = \max(\sigma_i^2)/\min(\sigma_i^2)$, ou seja, para $\sigma_i^2 = 1$ temos homoscedasticidade ($\lambda = 1$), caso contrário, para $\lambda > 1$ obtemos dados heteroscedásticos.

4.1.2 Regressão Logística

Modelos para dados binários são aqueles que apresentam resposta dicotômica assumindo apenas valores 0 ou 1. A regressão logística objetiva modelar o comportamento da variável binária condicionada a variáveis explicativas através de uma estrutura regressiva que permite estimar a probabilidade de ocorrência do evento de interesse. Esse modelo classifica-se como um caso particular dos MLGs sendo que a distribuição da variável de interesse (binomial) pertence à família de distribuições exponencial. Há diversas distribuições inclusas nessa classe que podem ser atribuídas a y_i em um MLG. Sua flexibilidade encontra-se na relação funcional entre a média de $\boldsymbol{y}$ e o preditor linear que pode ser arbitrada pela função de ligação, garantindo que $\mathbb{E}(\boldsymbol{y})$ como função de $\boldsymbol{\eta} = X\boldsymbol{\beta}$ não possua valores fora do seu espaço paramétrico (PAULA, 2013).

Seja Y uma variável aleatória com distribuição de Bernoulli onde p_i é a probabilidade de ocorrência de sucesso $\mathbb{P}(Y = 1)$ ou fracasso $\mathbb{P}(Y = 0)$. Sua função de probabilidade é dada por

$$f(y_i) = \mathbb{P}(Y = y_i) = p_i^{y_i}(1 - p_i)^{1-y_i}, \quad y_i = 0, 1, \quad 0 \leq p_i \leq 1.$$

A média e a variância de y_i são, respectivamente,

$$\mathbb{E}(y_i) = p_i,$$
$$\text{var}(y_i) = p_i(1 - p_i).$$

Os MLGs são constituídos pela componente aleatória, componente sistemática e a função de ligação $g(\cdot)$ monótona e diferenciável:

$$g(p_i) = \boldsymbol{\eta} = X\boldsymbol{\beta},$$

em que X é a matriz $n \times k$ de covariáveis, $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_k)^\top$ é o k-vetor de parâmetros, a componente sistemática é a soma linear dos efeitos $\boldsymbol{x}_i^\top \boldsymbol{\beta}$ das covariáveis e a função de ligação $g(p_i)$ relaciona a probabilidade de sucesso com o preditor linear. A regressão logística considera a função de ligação logit, expressa por

$$g(p_i) = \log\left(\frac{p_i}{1 - p_i}\right).$$

Ela restringe a média de y_i no intervalo unitário e relaciona aos preditores que variam no conjunto dos reais. Neste caso, a ligação logit ainda possui conveniências estatísticas[1], pois é uma função de ligação canônica onde o preditor linear modela diretamente o parâmetro canônico.

[1] Para mais detalhes, ver (PAULA, 2013, p. 8).

O modelo tem uma formulação equivalente a

$$p_i = \frac{1}{1+\exp(-\eta_i)}.$$

Analogamente, alguns autores escrevem

$$p_i = \frac{\exp(\eta_i)}{1+\exp(\eta_i)}.$$

Uma vantagem de ajustar uma regressão logística caracteriza-se pela interpretabilidade dos parâmetros baseada na razão de chances de probabilidades (*Odds Ratio* - OR). É uma medida importante para mensurar a força de associação entre as variáveis ao comparar a chance de sucesso de uma determinada característica p_i^* em relação a p_i. Obtemos OR, aplicando a função exponencial diretamente nos coeficientes do modelo, como segue,

$$\mathrm{OR}(x_1+1, x_1) = \frac{p_i^*/(1-p_i^*)}{p_i/(1-p_i)} = \exp(\widehat{\beta}_1),$$

isto é, a chance do evento ocorrer entre indivíduos que diferem em uma unidade na variável x_1 é igual a $\exp(\widehat{\beta}_1)$. Generalizando, para quantificar a alteração da chance de ocorrência de um evento após c incrementos unitários, utiliza-se $\exp(c\,\widehat{\beta}_1)$.

Os parâmetros desconhecidos $\boldsymbol{\beta}$ serão estimados através do popular método de máxima verossimilhança (PAWITAN, 2001). A função de verossimilhança baseada na distribuição de y_i é dada por

$$L(\boldsymbol{\beta}) = \prod_{i=1}^{n} p_i^{y_i}(1-p_i)^{1-y_i}.$$

Tomando o logaritmo, tem-se a função de log-verossimilhança:

$$\ell(\boldsymbol{\beta}) = \sum_{i=1}^{n} \big[y_i \log(p_i) + (1-y_i)\log(1-p_i)\big]. \tag{4.5}$$

Os estimadores de máxima verossimilhança são os valores que maximizam a função de verossimilhança, ou equivalentemente, a função de log-verossimilhança facilitando a otimização. Para obtenção dos estimadores, é preciso calcular o vetor escore (gradiente) e resolver $\boldsymbol{U}(\boldsymbol{\beta}) = \mathbf{0}$. Ou seja, deriva-se a log-verossimilhança em relação a cada parâmetro $\boldsymbol{\beta}$ e iguala-se a zero. O vetor escore para $\beta_j,\ j = 1,\ldots,k$ é dado por

$$\boldsymbol{U}(\boldsymbol{\beta}) = \frac{\partial \ell}{\partial \boldsymbol{\beta}} = \sum_{i=1}^{n} (y_i - p_i)\, x_i\,.$$

Contudo, o sistema $\boldsymbol{U}(\boldsymbol{\beta}) = \mathbf{0}$ não possui solução em forma fechada. Portanto, são considerados métodos numéricos como o algoritmo de otimização não linear quasi-Newton BFGS (NOCEDAL; WRIGHT, 1999) para maximização da função de log-verossimilhança.

4.2 MODELOS DIFERENCIALMENTE PRIVADOS

Há diversos métodos para alcançar privacidade em modelos matemáticos. Muitos têm se mostrado suscetíveis a ataques de inversão, principalmente quando o invasor possui algum conhecimento prévio sobre os indivíduos contidos no banco de dados (GANTA; KASIVISWANATHAN; SMITH, 2008). Como alternativa para impedir vazamentos de informações em problemas de regressão, DP caracteriza-se como uma metodologia adequada para preservar a privacidade dos indivíduos. Em geral, algoritmos que satisfazem as definições de DP são robustos a ataques. Dessa forma, um invasor não consegue inferir sobre as informações pessoais (contidas nas covariáveis) apenas acessando os resultados estimados e divulgados do modelo.

Neste sentido, modelos de regressão DP têm atuado fortemente na minimização de risco empírico (ERM - *Empirical Risk Minimization*) para seleção de modelos com menor erro possível. Estes incorporam abordagens de regularização para evitar problemas de sobreajuste (*overfitting*), principalmente quando os dados de treino e teste variam muito. O método é implementado adicionando um termo de penalidade que controla a função objetivo obtendo menor variância (CHAUDHURI; MONTELEONI, 2009; CHAUDHURI; MONTELEONI; SARWATE, 2011; KIFER; SMITH; THAKURTA, 2012).

Habitualmente, a aplicação de um algoritmo matemático requer a otimização de uma função objetivo de tal forma

$$J(\boldsymbol{\beta}) = \frac{1}{n}\sum_{i=1}^{n} l(x_i, y_i; \boldsymbol{\beta}),$$

em que $x_i = (x_{i1}, \ldots, x_{ik})^\top$ e $l(\cdot)$ é a função de custo. Os valores dos parâmetros $\boldsymbol{\beta}$ que otimizam $J(\boldsymbol{\beta})$ são os valores ótimos $\widehat{\boldsymbol{\beta}}$ resultantes da minimização ou maximização da função objetivo. Para o ajuste da regressão linear, estima-se os $\widehat{\boldsymbol{\beta}}^*$ que minimizam a soma dos quadrados de resíduos, tal que $\widehat{\boldsymbol{\beta}}^* = \arg\min_\beta \left[J(\boldsymbol{\beta})\right]$. Já no modelo logístico, maximizamos a função de log-verossimilhança.

Um modelo de regressão é definido diferencialmente privado ao incorporar uma perturbação para "mascarar" os dados sensíveis. Os métodos distinguem-se pelo tipo de perturbação utilizada sendo as principais classificadas por:

Perturbação na entrada (*input perturbation*): A intuitiva ideia de perturbar diretamente os dados verdadeiros foi formalizada por Traub, Yemini e Woźniakowski (1984). A perturbação é realizada ao adicionar ruído de forma aleatória em cada observação das covariáveis do banco de

dados, como segue

$$\boldsymbol{x_i}^{\text{pert}} = \boldsymbol{x_i} + \boldsymbol{z_i}, \tag{4.6}$$

em que "pert" abrevia perturbado, $\boldsymbol{x_i} = (x_{i1}, \ldots, x_{ik})^\top$ e $\boldsymbol{z_i} = (z_{i1}, \ldots, z_{ik})^\top$ é o ruído aleatório seguindo uma determinada função de probabilidade. Em geral, $\boldsymbol{z_i}$ é um vetor aleatório k-dimensional com densidade $\exp\left(-\frac{\varepsilon}{2}||\boldsymbol{z_i}||_2\right)$, em que $||\cdot||_2$ é a norma euclidiana (CHAUDHURI; MONTELEONI; SARWATE, 2011) ou cada elemento do vetor pode seguir uma lei de Laplace como a dada em (3.3), uma vez que ambos impõem um limite multiplicativo descrito pelo parâmetro de privacidade ε. Há métodos mais sofisticadas para escolha do ruído na literatura, a exemplo, Dwork *et al.* (2006) considera ruídos gaussianos e exponencialmente distribuídos para ocultar as respostas verdadeiras. A privacidade do modelo é preservada utilizando/divulgando os dados modificados x^{pert} (FUKUCHI; TRAN; SAKUMA, 2017; KANG *et al.*, 2020). Destaca-se que a teoria DP utilizando dados contínuos difere para dados discretos. Esta perturbação é aplicada exclusivamente considerando valores contínuos.

Perturbação na saída (*output perturbation*): é utilizada com mais frequência devido sua simplicidade de implementação. Consiste em inserir ruído pontualmente nos resultados após otimização do modelo, ou seja, as inferências são obtidas com os dados reais e apenas sua versão perturbada é disponibilizada. A perturbação é alcançada de tal forma:

$$\widehat{\boldsymbol{\beta}}^{\text{pert}} = \arg\min_{\beta} \left[J(\boldsymbol{\beta})\right] + \boldsymbol{z},$$

em que $\widehat{\boldsymbol{\beta}}^{\text{pert}}$ é um vetor k-dimensional contendo as estimativas dos parâmetros do modelo acrescida de ruído e $\boldsymbol{z}$ é um vetor aleatório k-dimensional com densidade $\exp\left(-\frac{\varepsilon}{\Delta f}||\boldsymbol{z}||_2\right)$, onde Δf é a sensibilidade da função de custo (CHAUDHURI; MONTELEONI; SARWATE, 2011). Eventualmente, se a coleta de dados for realizada via sistema de consulta (*queries*), o algoritmo – para preservar a privacidade – deve variar crescentemente a perturbação nas saídas conforme mais consultas forem solicitadas ao servidor (BECK, 1980).

Perturbação objetiva: Caracteriza-se como um dos algoritmos mais eficazes para modelagem diferencialmente privada (NEEL *et al.*, 2019). Neste caso, a adição de ruído é realizada diretamente na função de custo antecedendo a otimização, tal que

$$J_{\text{pert}}(\boldsymbol{\beta}) = J(\boldsymbol{\beta}) + \frac{1}{n}\boldsymbol{z}^\top\boldsymbol{\beta},$$

em que $\boldsymbol{z}$ é um vetor k-dimensional de ruído aleatório com distribuição definida como em (4.6). Assim, as estimativas diferencialmente privadas de $\boldsymbol{\beta}$ do modelo são obtidos otimizando a função

perturbada, tal que

$$\widehat{\boldsymbol{\beta}}^{\,\text{pert}} = \arg\min_{\beta} \left[J_{\text{pert}}(\boldsymbol{\beta}) \right],$$

ou

$$\widehat{\boldsymbol{\beta}}^{\,\text{pert}} = \arg\max_{\beta} \left[J_{\text{pert}}(\boldsymbol{\beta}) \right].$$

Trabalhos influentes na área de DP que aplicam este método são Chaudhuri, Monteleoni e Sarwate (2011); Zhang *et al.* (2012); Kifer, Smith e Thakurta (2012).

Perturbação no gradiente: A técnica mais recente de perturbação para garantia de privacidade baseia-se no método de gradiente descendente (BASSILY; SMITH; THAKURTA, 2014). O esquema iterativo é dado por

$$\beta_{t+1} = \beta_t - \zeta \left[\nabla J(\beta_t) + z_t \right],$$

em que ∇ é o gradiente, ζ é a taxa de aprendizado e z_t segue uma distribuição normal incorporando um ruído gaussiano. O processo inicia na iteração $t = 0$ e finaliza quando encontra o valor ótimo (e diferencialmente privado) de β.

Basicamente todas as técnicas envolvem adição de ruído alternando a etapa de aplicação no processo de modelagem. A perturbação baseada no gradiente apresenta comportamento insatisfatório (SONG; CHAUDHURI; SARWATE, 2013), porém métodos aperfeiçoados estão sendo desenvolvidos no estado da arte (WU *et al.*, 2017). A perturbação imposta diretamente nos dados reais (entrada) é viável quando há uma conexão dos dados com um servidor/agregador e objetiva-se proteger essa "comunicação" tornando os indivíduos não-identificáveis ao coletor de dados (FUKUCHI; TRAN; SAKUMA, 2017; DUCHI; JORDAN; WAINWRIGHT, 2013). Caso contrário, sua aplicação pode prejudicar excessivamente a acurácia do modelo pois a garantia de privacidade requer uma acentuada contaminação dos dados.

4.2.1 Proposta

Tratando-se de preservação de privacidade em regressão linear e logística, usualmente adota-se as abordagens de perturbação na saída e objetiva embasadas pela maior adequabilidade da calibração de ruído e devido as adversidades apontadas das demais técnicas de perturbação. Portanto, o presente trabalho considera a perturbação objetiva, a qual, mostra-se uma abordagem superior na literatura de regressão DP comparativamente à perturbação na saída (CHAUDHURI; MONTELEONI; SARWATE, 2011). Salientamos que a proposta baseia-se na metodologia do mecanismo de Laplace (Seção 3.6.3) para incorporar a aleatoriedade no modelo. Consequentemente, os possíveis resultados que a saída do algoritmo pode assumir são restritos ao espaço

Euclidiano e, portanto, utilizamos a norma-ℓ_2 como métrica para calibrar a sensibilidade das funções.

No modelo de regressão linear, alcançamos privacidade minimizando a função objetivo perturbada que neste caso é a soma dos quadrados dos resíduos (4.1) com penalização de ruído. Os parâmetros estimados sob DP são dados por

$$\widehat{\boldsymbol{\beta}}^{*} = \arg\min_{\beta} \left\{ \frac{1}{n} \left[(\boldsymbol{y} - X\boldsymbol{\beta})^{\top} (\boldsymbol{y} - X\boldsymbol{\beta}) + \boldsymbol{z}^{\top}\boldsymbol{\beta} \right] \right\},$$

em que $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_k)^{\top}$, X é a matriz $n \times k$ de covariáveis e $\boldsymbol{z}$ é um vetor aleatório k-dimensional de distribuição $\exp\left(-\frac{\varepsilon}{2}||\boldsymbol{z}||_2\right)$ com valores de ε baseados no Algoritmo 2 de (CHAUDHURI; MONTELEONI; SARWATE, 2011, p. 1077).

Na modelagem de regressão logística, a função de custo é definida pela função de log-verossimilhança (4.5) penalizada com ruído. As estimativas diferencialmente privadas dos parâmetros são obtidas por

$$\widehat{\boldsymbol{\beta}}^{*} = \arg\max_{\beta} \left\{ \frac{1}{n} \sum_{i=1}^{n} \left[y_i \log(p_i) + (1 - y_i) \log(1 - p_i) \right] + \frac{1}{n} \boldsymbol{z}^{\top}\boldsymbol{\beta} \right\},$$

em que $\boldsymbol{z}$ é um vetor k-dimensional de ruído aleatório composto por ocorrências de uma distribuição gama[2] com suporte restrito nos reais positivos conforme proposto em Chaudhuri e Monteleoni (2009). Os estimadores baseados nessa função serão denominados MVDP (máxima verossimilhança diferencialmente privada) exclusivamente neste trabalho.

4.2.2 Modelos privados usuais

Pelo propósito comparativo, consideramos na aplicação prática (Seção 5.2) outras duas abordagens usuais na literatura de DP para regressão linear e logística. A primeira é o *Functional Mechanism* - FM (ZHANG *et al.*, 2012), o qual consiste na perturbação objetiva de uma função para alcançar privacidade — similar a proposta deste trabalho — porém a adição de ruído Laplaciano é restrita aos primeiros termos polinomiais da função objetivo particionada por uma aproximação de segunda ordem, tal que para cada β_j, $j = 1, \ldots, k$, temos

$$J_{\text{pert}}(\beta_j) = \left[\mathfrak{a} + \text{Lap}(\Delta f / \varepsilon) \right] {\beta_j}^2 + \left[\mathfrak{b} + \text{Lap}(\Delta f / \varepsilon) \right] \beta_j,$$

em que $\mathfrak{a}$ e $\mathfrak{b}$ representam os coeficientes da função. Na regressão linear, a função objetivo é definida por

[2] A distribuição de Laplace caracteriza-se como um caso particular da distribuição gama.

$$J(\boldsymbol{\beta}) = \sum_{i=1}^{n} \left(y_i - \boldsymbol{x}_i^\top \boldsymbol{\beta} \right)^2,$$

e os seus coeficientes são dados, respectivamente, por

$$\mathfrak{a} = \sum_{1 \leq j,d \leq k} \left(\sum_{i=1}^{n} x_{ij} x_{id} \right), \qquad \mathfrak{b} = \sum_{j=1}^{k} \left(2 \sum_{i=1}^{n} y_i x_{ij} \right), \tag{4.7}$$

em que k é o número de variáveis regressoras (covariáveis). A sensibilidade Δf da função objetivo linear é igual a $2(k+1)^2$, ou seja, o algoritmo adiciona ruído aleatório em cada coeficiente gerado da distribuição $\text{Lap}\big(2(k+1)^2/\varepsilon\big)$.

Na regressão logística, a função objetivo é dada por

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

Os termos da função não possuem solução em forma fechada e, consequentemente, não é possível expressá-la como polinômios de ordem finita. Alternativamente, os autores utilizam uma aproximação polinomial via expansão da série de Taylor para os dois termos da função, dada por

$$J(\boldsymbol{\beta}) = \sum_{i=1}^{n} \sum_{v=0}^{2} \frac{f_1^{(v)}(0)}{v!} \left(\boldsymbol{x}_i^\top \boldsymbol{\beta} \right)^v - \left(\sum_{i=1}^{n} y_i \boldsymbol{x}_i^\top \right) \boldsymbol{\beta},$$

em que a abordagem requer as derivadas $f_1^{(v)}(0)$ para $v = 0, 1, 2$ iguais a $\log 2$, $1/2$ e $1/4$, respectivamente. Tal aproximação da função objetivo logística resulta nos seguintes coeficientes:

$$\mathfrak{a} = \frac{1}{8} \sum_{i=1}^{n} \left(\boldsymbol{x}_i^\top \right)^2 \qquad \text{e} \qquad \mathfrak{b} = \frac{1}{2} \sum_{i=1}^{n} \boldsymbol{x}_i^\top - \sum_{i=1}^{n} y_i \boldsymbol{x}_i^\top. \tag{4.8}$$

Para modelar a regressão logística perturbada/privada, a quantidade do ruído de Laplace inserido nos monômios da função objetivo pré-otimização é calibrada pela sensibilidade $\Delta f = k^2/4 + 3k$.

A segunda abordagem usual (FANG *et al.*, 2019) estende o método FM no intuito de fornecer mais flexibilidade e aprimorar a acurácia do modelo. O trabalho problematiza o mecanismo precursor pela adição excessiva de ruído desconsiderando as distintas influências de cada monômio na função objetivo e propõe uma ponderação do ruído de acordo com a sensibilidade individual. Denominado *Differentiated Privacy Budget Allocation* (DPBA), o algoritmo utiliza a propriedade de composição diferencialmente privada para alocar diferentes valores de ε determinados pela sensibilidade de cada termo polinomial da função objetivo decomposta. Assim, a função objetivo perturbada é da forma

$$J_{\text{pert}}(\boldsymbol{\beta}) = \sum_{j=1}^{k} \left\{ \left[\mathfrak{a} + \text{Lap}(\Delta f_1/\varepsilon_1) \right] {\beta_j}^2 + \left[\mathfrak{b} + \text{Lap}(\Delta f_2/\varepsilon_2) \right] \beta_j \right\},$$

em que para ajustar o modelo de regressão linear diferencialmente privado, os coeficientes $\mathfrak{a}$ e $\mathfrak{b}$ equivalem-se a (4.7) e suas respectivas sensibilidades são iguais a $\Delta f_1 = 2k^2$ e $\Delta f_2 = 4k$. Para aplicação do DPBA ao modelar a regressão logística, $\mathfrak{a}$ e $\mathfrak{b}$ são definidos como na Equação (4.8) em que suas sensibilidades Δf_1 e Δf_2 são dadas respectivamente por $k^2/4$ e $3k$.

A metodologia utilizada para escolher $\Delta f_1, \Delta f_2, \varepsilon_1$ e ε_2 pode ser encontrada em (FANG *et al.*, 2019, p. 5). No geral, atribui-se maiores valores ao parâmetro de perda de privacidade ε para itens mais sensíveis, os quais possuem maior impacto no modelo.

5 AVALIAÇÃO NUMÉRICA

Um estudo de simulações de Monte Carlo é apresentado nesta seção para avaliar numericamente os estimadores diferencialmente privados dos parâmetros dos modelos de regressão linear e logístico via perturbação objetiva variando ε no intervalo $[0.01, 5]$. Foram consideradas 10000 réplicas de Monte Carlo em cada cenário variando os tamanhos amostrais tal que $n \in \{50, 100, 300, 1000\}$. As medidas estatísticas calculadas foram: média, viés, viés relativo percentual (VR%) e erro quadrático médio (EQM). As implementações computacionais foram desenvolvidas em linguagem R (R Core Team, 2020). Destaca-se que a adição de ruído, dependendo dos valores aleatórios gerados, pode acarretar em uma função sem solução ótima ou encontrar mínimos/máximos locais. Portanto, nos procedimentos de otimização das funções objetivo foi utilizado o método quasi-Newton BFGS (NOCEDAL; WRIGHT, 1999), o qual produziu melhor estabilidade para convergência das funções perturbadas.

No caso linear, investigamos a influência de DP aplicada em modelos de regressão com dados sob presença de heteroscedasticidade através da avaliação dos estimadores privados e consistentes. Portanto, impomos a função cedástica (4.4) no modelo tal que $\sigma_i^2 = \exp\{\gamma_1 + \gamma_2 w_i\}$. O parâmetro de dispersão $\boldsymbol{\gamma}$ foi definido igual a 0.25 tal que o grau de heteroscedasticidade seja aproximadamente $\lambda \approx 15$. Os valores da covariável w_i relacionada a $\boldsymbol{\gamma}$ foram gerados a partir da distribuição log-normal padrão. Nas simulações, definimos o tamanho amostral de w_i em $n = 50$ e a replicamos de acordo com os n maiores para garantir que o grau de heteroscedasticidade permaneça constante enquanto aumentamos o número de observações. Ou seja, cada valor de w_i foi replicado duas vezes quando $n = 100$, três vezes para $n = 150$, quatro vezes no caso de $n = 200$ e assim por diante. Essa técnica é usualmente empregada na literatura (CRIBARI-NETO; ZARKOS, 2001; MACKINNON; WHITE, 1985)

A Tabela 3 apresenta resultados numéricos referentes às estimativas do parâmetro β_2 do modelo de regressão linear DP estruturado por $y_i = \beta_1 + \beta_2 x_i + \sigma_i u_i,\ i = 1, \ldots, n$. A geração dos dados foi realizada utilizando $\beta_1 = 1$ e $\beta_2 = 3$. Os erros são independentes e identicamente distribuídos de acordo com a distribuição normal padrão $\mathcal{N}(0, 1)$. A covariável x_i foi obtida gerando aleatoriamente observações da distribuição uniforme igualmente espaçada no intervalo $[0, 1]$. Ao analisar os resultados, observa-se decréscimo do viés conforme o tamanho amostral aumenta, ou seja, a estimativa do parâmetro se aproxima do valor verdadeiro. A presença de heteroscedasticidade ocasionou em maior variabilidade visto que os estimadores considerando $\lambda \approx 15$ apresentaram maior EQM quando comparados ao caso homoscedástico ($\lambda = 1$). No

que se refere à privacidade, como esperado, para restrições fortes (ε menores) temos maiores distorções inferenciais. Essa distorção é intensificada para menores amostras e amenizada para maiores, ou seja, a perturbação causada por ruídos discrepantes (mais distantes de zero) tem maior influência para n inferior. Por exemplo, com $\varepsilon = 0.01$ e $n = 50$, obteve-se VR$= -2.23\%$, enquanto para $n = 1000$, o VR foi igual a -0.126% com estimativa média diferendo apenas 0.004 do parâmetro verdadeiro. Tal comportamento indica que, para preservar a privacidade, a adição de ruído deve ser ponderada com base no tamanho amostral e na escala de acurácia calibrada por ε.

Na Tabela 4 encontram-se os resultados das simulações de Monte Carlo para as estimativas $\widehat{\beta}_2$ do modelo de regressão logística estruturado por $\log\left(\frac{p_i}{1-p_i}\right) = \beta_1 + \beta_2 x_i$. A geração dos dados foi realizada com parâmetros definidos por $\beta_1 = 0$ e $\beta_2 = 2$. Os valores da covariável associada a β_2 foram obtidos a partir da geração aleatória de uma distribuição normal padrão. Verifica-se que as estimativas dos parâmetros melhoram à medida que o tamanho da amostra cresce, como é o caso do viés reduzido assintoticamente. Contudo, o maior interesse está em avaliar o comportamento sob diferentes níveis de privacidade. Nota-se que quanto menor for o valor de ε (forte privacidade), mais as estimativas se afastam do verdadeiro valor do parâmetro especificado. Por exemplo, para $n = 50$, o estimador $\widehat{\beta}_2$ possui VR igual a -9.80% com $\varepsilon = 0.01$ enquanto para $\varepsilon = 1$ é reduzido a -0.78%. De fato, a relação "acurácia $\times$ perda de privacidade" é verificada nestes modelos diferencialmente privados visto que os resultados apresentam-se mais distorcidos para maiores perturbações na função de log-verossimilhança (alta adição de ruído).

5.1 TAXAS DE REJEIÇÃO

Foram realizadas simulações de tamanho e poder para os testes z e quasi-t associados aos estimadores para mensurar o impacto da restrição de privacidade no desempenho dos modelos avaliados anteriormente considerando as mesmas configurações. Adicionalmente, foi incluído um cenário com grau de heteroscedasticidade acentuado onde $\lambda \approx 40$ para valores de dispersão $\boldsymbol{\gamma} = 0.42$. Vale salientar que testes desempenham bem para níveis de significância estimados (tamanho) próximos do nível nominal arbitrado e detecção apurada de modelos mal especificados (poder alto). Todas simulações foram realizadas considerando 10000 réplicas de Monte Carlo com tamanhos amostrais $n \in \{50, 100, 150, 200\}$. Testamos as hipóteses considerando níveis nominais de 5% e 10%.

Tabela 3 – Resultados da simulação de Monte Carlo para $\beta_2 = 3$ do modelo linear.

λ	ε	0.01	0.05	0.1	0.3	0.6	1	5
				$n = 50$				
1	Média	2.933	2.934	2.937	2.962	2.990	2.995	2.995
	Viés	−0.067	−0.066	−0.063	−0.038	−0.010	−0.005	−0.005
	VR (%)	**−2.232**	−2.198	−2.095	−1.272	−0.328	−0.156	**−0.154**
	EQM	0.260	0.260	0.259	0.257	0.255	0.255	0.255
15	Média	2.928	2.929	2.932	2.957	2.985	2.990	2.990
	Viés	−0.072	−0.071	−0.068	−0.043	−0.015	−0.010	−0.010
	VR (%)	−2.400	−2.366	−2.263	−1.440	−0.496	−0.325	−0.323
	EQM	0.592	0.592	0.591	0.588	0.587	0.587	0.587
				$n = 100$				
1	Média	2.968	2.969	2.971	2.985	2.998	2.999	2.999
	Viés	−0.032	−0.031	−0.029	−0.015	−0.002	−0.001	−0.001
	VR (%)	−1.057	−1.036	−0.971	−0.496	−0.081	−0.040	−0.040
	EQM	0.123	0.123	0.123	0.122	0.122	0.122	0.122
15	Média	2.966	2.966	2.968	2.983	2.995	2.996	2.996
	Viés	−0.034	−0.034	−0.032	−0.017	−0.005	−0.004	−0.004
	VR (%)	−1.142	−1.120	−1.055	−0.579	−0.163	−0.122	−0.122
	EQM	0.292	0.292	0.292	0.291	0.291	0.291	0.291
				$n = 300$				
1	Média	2.989	2.989	2.990	2.996	2.999	2.999	2.999
	Viés	−0.011	−0.011	−0.010	−0.004	−0.001	−0.001	−0.001
	VR (%)	−0.370	−0.358	−0.324	−0.122	−0.037	−0.035	−0.035
	EQM	0.040	0.040	0.040	0.040	0.040	0.040	0.040
15	Média	2.988	2.988	2.989	2.995	2.998	2.998	2.998
	Viés	−0.012	−0.012	−0.011	−0.005	−0.002	−0.002	−0.002
	VR (%)	−0.409	−0.398	−0.363	−0.162	−0.076	−0.075	−0.075
	EQM	0.218	0.218	0.218	0.218	0.217	0.217	0.217
				$n = 1000$				
1	Média	2.996	2.996	2.997	2.999	2.999	2.999	2.999
	Viés	−0.004	−0.004	−0.003	−0.001	−0.001	−0.001	−0.001
	VR (%)	**−0.126**	−0.120	−0.102	−0.035	−0.026	−0.026	−0.026
	EQM	0.012	0.012	0.012	0.012	0.012	0.012	0.012
15	Média	2.997	2.997	2.997	2.999	3.000	3.000	3.000
	Viés	−0.003	−0.003	−0.003	−0.001	0.000	0.000	0.000
	VR (%)	−0.109	−0.103	−0.086	−0.018	−0.010	−0.010	−0.010
	EQM	0.012	0.012	0.012	0.012	0.012	0.012	0.012

O interesse reside em testar se um valor hipotético equivale ao verdadeiro valor do parâmetro, isto é, a hipótese nula $\mathcal{H}_0 : \beta_j = \beta_j^0$ contra a hipótese alternativa $\mathcal{H}_1 : \beta_j \neq \beta_j^0$ em

Tabela 4 – Resultados da simulação de Monte Carlo para $\beta_2 = 2$ do modelo logístico.

ε	0.01	0.05	0.1	0.3	0.6	1
			$n = 50$			
Média	1.804	1.804	1.806	1.825	1.882	1.984
Viés	−0.196	−0.196	−0.194	−0.175	−0.118	−0.016
VR (%)	**−9.804**	−9.775	−9.686	−8.760	−5.923	**−0.779**
EQM	0.276	0.276	0.276	0.277	0.288	0.335
			$n = 100$			
Média	1.907	1.907	1.908	1.918	1.947	1.997
Viés	−0.093	−0.093	−0.092	−0.082	−0.053	−0.003
VR (%)	−4.658	−4.643	−4.596	−4.111	−2.655	−0.125
EQM	0.157	0.157	0.157	0.158	0.162	0.175
			$n = 300$			
Média	1.968	1.969	1.969	1.972	1.982	1.998
Viés	−0.032	−0.031	−0.031	−0.028	−0.018	−0.002
VR (%)	−1.576	−1.571	−1.555	−1.391	−0.904	−0.079
EQM	0.053	0.053	0.053	0.053	0.053	0.054
			$n = 1000$			
Média	1.990	1.990	1.990	1.991	1.994	1.999
Viés	−0.010	−0.010	−0.010	−0.009	−0.006	−0.001
VR (%)	**−0.484**	−0.482	−0.478	−0.428	−0.282	−0.036
EQM	0.017	0.017	0.017	0.017	0.017	0.017

que β_j^0 é uma constante para $j = 1, 2$. A estatística de teste é dada por

$$\tau = \frac{\widehat{\beta}_j - \beta_j^0}{\widehat{\text{ep}}(\widehat{\beta}_j)},$$

em que $\widehat{\text{ep}}$ denota a estimativa do erro padrão de $\widehat{\beta}_j$ representada por $\sqrt{\widehat{\text{var}}(\widehat{\beta}_j)}$. A variância estimada de $\widehat{\beta}_j$ é obtida a partir do estimador consistente da matriz de covariância. Sob $\mathcal{H}_0$, a estatística de teste possui distribuição assintótica normal padrão. Rejeitamos a hipótese nula ao nível de significância nominal α se o valor absoluto de τ for maior que o valor crítico dado pelo quantil $1 - \alpha/2$ da distribuição $\mathcal{N}(0, 1)$.

5.1.1 Tamanho do teste

A Tabela 5 apresenta as taxas de rejeição nulas dos testes quasi-t para testar $\mathcal{H}_0 : \beta_2 = 3$ do modelo de regressão linear $y_i = \beta_1 + \beta_2 x_i + \sigma_i u_i$. O resultado das simulações determina a proporção de vezes que a hipótese nula foi incorretamente rejeitada. Quanto mais próximo esse nível de significância estimado for do nominal, melhor desempenho possui o

teste. Como esperado, o estimador MQO apresenta maiores distorções de tamanho sob heteroscedasticidade comparativamente a quando este cenário não se verifica. Por exemplo, tomando como base o nível nominal 5% para $n = 100$ e $\varepsilon = 0.01$, na presença de erros homoscedásticos o teste baseado no MQO rejeita $\mathcal{H}_0$ 5.34% das vezes, enquanto quando há desvio de homoscedasticidade ($\lambda \approx 15$) essa taxa cresce para 7.09% e distancia-se para 9.57% ao aumentar o grau da heteroscedasticidade. Com exceção do HC0 em amostras pequenas ($n = 50$), os estimadores consistentes apresentam desempenho superior ao MQO sob heteroscedasticidade com destaque para o HC4 que viabiliza testes mais confiáveis. Ainda, observamos que fortes restrições de privacidade (ε pequeno) conduzem a testes cujas taxas de rejeição encontram-se mais distantes do nível nominal especificado.

Sob suspeita de heteroscedasticidade comumente utiliza-se estimadores consistentes para estimar os erros-padrão e inferir sobre os coeficientes de um modelo. Porém, em modelos não-lineares, particularmente na regressão logística, não é usual considerar uma estrutura cedástica. Caso os erros fossem heteroscedásticos, a estimação dos parâmetros e da matriz de covariância via máxima verossimilhança seria viesada e inconsistente; a não ser que a função seja adaptada para considerar a presença de heteroscedasticidade (GREENE, 2003, p. 673–674).

Na Tabela 6 apresentamos as taxas de rejeição do teste z sob a hipótese nula baseado no estimador MVDP para testar $\mathcal{H}_0 : \beta_2 = 3$ do modelo logístico $\log\left(\frac{p_i}{1-p_i}\right) = \beta_1 + \beta_2 x_i$ variando o nível de significância $\alpha \in \{5\%, 10\%\}$. Tratando-se da regressão logística, desconsideramos os estimadores HC0 e HC4 pois não modelamos uma função cedástica. De acordo com os resultados podemos notar que conforme o tamanho amostral aumenta os tamanhos do teste se aproximam dos níveis nominais especificados. As distorções de tamanho tendem a diminuir quando reduzimos a privacidade (aumentando ε). Por exemplo em 6a, no caso de $n = 50$ percebe-se inicialmente uma taxa de rejeição que ultrapassa 8% com $\varepsilon = 0.01$, enquanto para $\varepsilon = 5$, o teste é menos liberal com taxa 5.59%.

5.1.2 Poder do teste

Esta seção objetiva medir o poder de detecção dos testes em modelos mal especificados, ou seja, a capacidade de identificar que a hipótese nula é falsa quando de fato ela é. Consideramos os mesmos cenários utilizados nas simulações de tamanho, contudo, agora a geração de dados é realizada impondo a hipótese alternativa.

Tabela 5 – Taxas de rejeição nula (tamanho do teste) do modelo linear.

n	ε	0.01		0.1		0.5		1		5	
	α	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%
						$\lambda = 1$					
	MQO	5.87	10.93	5.64	10.95	5.51	10.81	5.76	10.48	5.31	10.58
50	HC0	7.31	12.71	6.98	12.53	6.70	12.13	6.81	12.08	6.60	12.25
	HC4	5.99	10.94	5.63	10.96	5.60	10.73	5.77	10.34	5.57	10.74
	MQO	**5.34**	10.43	5.20	10.45	5.25	10.44	5.45	10.19	5.07	10.14
100	HC0	5.94	11.14	5.65	11.17	6.04	11.30	6.04	10.84	5.58	10.95
	HC4	5.42	10.43	5.24	10.52	5.31	10.48	5.43	10.11	5.16	10.16
	MQO	5.37	10.34	5.03	10.34	5.19	10.37	5.20	10.18	5.24	10.25
150	HC0	5.87	10.78	5.37	10.91	5.46	10.70	5.56	10.63	5.84	10.82
	HC4	5.46	10.33	5.01	10.29	5.12	10.27	5.28	10.21	5.50	10.29
	MQO	5.08	10.12	5.37	10.31	5.07	10.28	5.20	10.01	4.95	10.10
200	HC0	5.46	10.40	5.56	10.82	5.51	10.76	5.56	10.44	5.27	10.37
	HC4	5.16	10.01	5.39	10.34	5.17	10.20	5.25	10.00	5.05	10.01
						$\lambda \approx 15$					
	MQO	4.66	9.12	4.65	9.13	5.24	8.44	8.36	14.83	10.11	17.07
50	HC0	6.80	12.26	6.80	12.22	6.82	11.62	6.50	12.24	6.82	12.51
	HC4	5.76	10.68	5.76	10.62	5.49	10.50	5.30	10.45	5.49	10.60
	MQO	**7.09**	12.51	7.10	12.51	7.71	17.79	5.86	11.7	3.76	8.12
100	HC0	5.98	10.99	5.96	10.94	5.94	11.24	5.52	10.95	5.67	10.84
	HC4	5.46	10.28	5.46	10.29	5.37	10.09	5.07	10.16	5.11	10.10
	MQO	4.40	8.85	4.41	8.85	5.92	10.43	7.55	13.66	2.835	6.32
150	HC0	5.71	10.77	5.70	10.75	5.41	11.40	5.30	10.68	5.55	10.42
	HC4	5.31	10.27	5.30	10.29	5.00	10.72	5.04	10.09	5.12	9.88
	MQO	8.44	14.98	8.40	14.94	2.99	8.71	3.85	8.075	2.99	6.97
200	HC0	5.49	10.76	5.47	10.77	5.18	10.38	5.37	10.45	5.29	10.54
	HC4	5.08	10.29	5.07	10.29	4.91	10.01	5.10	10.06	5.05	10.11
						$\lambda \approx 40$					
	MQO	8.31	15.38	11.88	8.22	4.23	7.94	5.97	6.87	6.90	7.50
50	HC0	6.10	13.08	5.95	12.28	5.96	12.30	5.51	11.58	6.31	11.62
	HC4	4.54	10.62	4.55	10.26	5.14	11.14	4.58	10.78	5.12	9.62
	MQO	**9.57**	11.38	5.77	13.28	3.28	13.10	4.58	12.70	3.13	12.50
100	HC0	5.36	11.42	5.42	11.30	5.29	11.31	5.57	10.98	5.28	11.02
	HC4	4.86	9.54	4.91	10.54	4.94	10.38	5.15	10.30	4.85	9.48
	MQO	4.70	8.10	8.60	14.38	11.42	14.36	3.08	17.02	5.27	8.94
150	HC0	5.07	10.70	5.30	10.86	5.24	10.70	5.24	10.40	5.29	10.33
	HC4	5.14	10.14	4.90	10.14	4.92	10.24	4.97	10.84	5.04	10.98
	MQO	4.05	14.16	5.23	9.92	7.31	8.34	2.98	8.04	6.32	11.09
200	HC0	5.29	10.64	5.21	10.58	5.65	10.52	5.39	9.88	5.46	10.09
	HC4	5.07	10.21	4.99	10.12	5.22	10.12	5.15	10.32	5.23	9.78

Os resultados das simulações de Monte Carlo para avaliar o poder do teste nos modelos privados de regressão linear e logístico estão presentes, respectivamente, nas Tabelas 7 e 8. No modelo linear, testamos $\mathcal{H}_0 : \beta_2 = 4$ sendo o verdadeiro valor do parâmetro igual a 5.

Tabela 6 – Taxas de rejeição nulas (tamanho do teste) do modelo logístico.

(a) $\alpha = 5\%$

ε \ n	50	100	150	200
0.01	**8.01**	6.25	5.47	4.98
0.1	7.14	5.96	5.33	5.34
0.5	6.46	5.33	5.09	5.12
1	5.70	4.75	5.19	4.71
5	**5.59**	4.89	4.84	4.84

(b) $\alpha = 10\%$

ε \ n	50	100	150	200
0.01	12.56	10.17	9.78	10.15
0.1	11.35	10.39	10.28	10.37
0.5	11.04	9.63	9.56	9.85
1	9.25	9.84	10.10	10.01
5	9.21	9.10	9.38	9.80

Observa-se que as taxas de rejeição não nula (expressas em porcentagem) crescem conforme a amostra aumenta. Ainda, no caso de homoscedasticidade, os poderes dos testes são similares entre os estimadores considerados. Já perante heteroscedasticidade, os estimadores consistentes HC0 e HC4 desempenham testes mais poderosos quando comparados ao estimador usual MQO.

Na Tabela 8, a geração dos dados foi realizada utilizando $\beta_2 = 2$ (verdadeiro valor do parâmetro), porém testamos $\mathcal{H}_0 : \beta_2 = 1$ na intenção de verificar a capacidade do teste para detectar a incorreta especificação do modelo. Nota-se que os testes aproximam-se de 100% conforme o tamanho amostral aumenta, ou seja, tornam-se mais poderosos. Em geral, a obtenção de um modelo mais privado (ε próximo de zero) implica em menores poderes devido a presença de maior aleatoriedade no processo de estimação. O ruído aleatório dificulta a detecção do teste fazendo com que aumente a probabilidade de não rejeitar $\mathcal{H}_0$ quando $\mathcal{H}_0$ é falsa.

5.2 APLICAÇÃO EMPÍRICA

Esta seção objetiva demonstrar a aplicabilidade prática dos modelos diferencialmente privados ajustando dados reais e avaliar a influência privacidade no desempenho destes. Comparamos o modelo proposto com sua versão não-privada e com alternativas presentes na literatura de regressão e privacidade. São elas FM e DPBA (abordadas na Seção 4.2.1). Foi ajustado um modelo de regressão linear diferencialmente privado e para a aplicação da regressão logística consideramos dois bancos de dados reais. É importante salientar que as simulações de diagnóstico com os dados reais (exceto os resultados da Tabela 9) foram repetidas 100 vezes: devido ao ruído aleatório inserido no processo de estimação é comum que haja uma variação nos coeficientes de cada novo modelo ajustado. Portanto, na intenção de avaliar mais precisamente o comportamento do mecanismo DP apresentamos a média desses resultados estimados sob as 100 réplicas.

Tabela 7 – Taxas de rejeição não-nulas (poder do teste) do modelo linear.

n	ε	0.01		0.1		0.5		1		5	
	α	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%
						$\lambda = 1$					
50	MQO	45.69	57.56	31.99	59.72	51.89	66.19	56.93	63.14	50.95	65.85
	HC0	48.40	59.56	35.61	62.00	54.34	68.08	58.94	64.98	53.40	67.89
	HC4	44.77	56.32	30.39	59.13	50.92	65.57	56.07	62.28	50.03	65.11
100	MQO	79.43	84.35	84.25	90.89	84.76	86.44	85.65	86.14	78.05	89.79
	HC0	80.14	84.68	84.79	91.11	85.14	86.72	86.14	86.69	78.79	90.22
	HC4	78.92	83.70	83.82	90.54	84.26	85.46	85.34	85.90	77.63	89.61
150	MQO	**92.12**	95.84	93.62	97.70	89.72	96.27	94.69	94.60	**94.71**	97.28
	HC0	92.41	95.99	93.77	97.78	90.03	96.39	94.90	94.71	94.81	97.34
	HC4	91.97	95.69	93.45	97.66	89.23	96.19	94.64	94.44	94.48	97.23
200	MQO	98.50	97.98	97.83	99.56	95.87	98.91	97.08	98.61	98.75	99.55
	HC0	98.54	97.98	97.91	99.57	96.03	98.89	97.13	98.67	98.79	99.57
	HC4	98.48	97.86	97.81	99.54	95.81	98.84	97.04	98.58	98.76	99.55
						$\lambda \approx 15$					
50	MQO	22.93	33.62	25.64	34.38	25.68	38.22	25.94	43.82	25.43	40.27
	HC0	28.21	38.70	30.70	38.66	30.05	45.15	33.90	37.88	31.02	47.62
	HC4	25.23	36.00	27.77	35.83	26.88	42.97	30.90	34.97	28.02	45.33
100	MQO	45.07	51.85	40.32	62.92	46.31	57.99	53.40	55.95	39.16	58.99
	HC0	48.60	48.68	41.24	59.44	51.43	49.49	59.85	60.64	46.23	64.01
	HC4	47.24	47.54	38.87	57.99	50.17	47.45	58.15	59.21	44.20	62.79
150	MQO	**65.81**	69.97	57.24	54.45	69.11	74.87	58.04	73.23	59.91	75.70
	HC0	**73.81**	73.28	59.88	53.69	71.46	75.78	54.65	80.92	53.79	74.13
	HC4	**73.03**	72.53	58.94	52.52	70.50	75.14	53.53	80.00	52.70	73.57
200	MQO	75.25	69.96	68.95	78.08	72.68	83.17	76.13	85.38	57.51	90.97
	HC0	79.26	64.10	68.57	83.98	70.19	85.08	75.61	89.59	53.61	92.20
	HC4	78.68	63.25	67.68	83.42	69.58	84.71	75.03	89.27	52.14	91.98
						$\lambda \approx 40$					
50	MQO	16.92	25.25	15.53	26.32	12.12	22.14	15.65	20.87	17.89	22.31
	HC0	16.22	24.49	20.63	26.89	22.16	31.69	20.04	32.70	26.10	31.08
	HC4	13.48	21.65	18.88	25.03	19.57	29.37	17.65	30.42	24.10	28.55
100	MQO	23.88	30.27	22.93	44.43	19.36	35.14	26.97	45.84	31.20	37.04
	HC0	27.90	33.88	22.38	47.94	33.48	48.43	28.72	47.38	38.13	46.86
	HC4	26.26	32.54	20.41	46.58	31.88	47.22	27.15	45.96	36.69	45.39
150	MQO	39.82	39.95	34.01	65.49	45.08	52.59	29.89	48.10	41.80	61.41
	HC0	42.05	40.37	34.73	74.87	49.28	45.78	51.52	50.94	51.96	64.57
	HC4	41.16	39.69	33.95	74.20	48.41	45.02	50.48	50.13	50.88	63.97
200	MQO	41.64	67.13	45.71	51.55	52.41	54.65	49.76	64.46	50.34	67.85
	HC0	57.24	65.93	58.00	59.59	60.07	60.47	54.36	58.75	67.96	68.85
	HC4	56.46	65.34	57.37	58.98	59.26	59.91	53.54	58.26	67.10	68.24

5.2.1 Regressão Linear

Dados médicos: Obtidos do livro de Lantz (2019), o conjunto de dados contendo 1338 observações visa modelar as despesas médicas dos cidadãos explicada por covariáveis de cunho pessoal como idade, sexo, índice de massa corporal, número de filhos e se o indivíduo é fumante ou não.

Tabela 8 – Taxas de rejeição não-nulas (poder do teste) modelo logístico.

(a) $\alpha = 5\%$

ε \ n	50	100	150	200
0.01	14.01	68.88	94.06	96.66
0.1	15.86	74.62	88.26	98.30
0.5	23.56	74.80	85.60	98.43
1	40.59	71.34	93.50	97.70
5	33.24	80.57	96.44	98.49

(b) $\alpha = 10\%$

ε \ n	50	100	150	200
0.01	38.74	83.58	97.74	98.72
0.1	38.94	84.66	95.86	99.32
0.5	47.48	87.25	93.23	99.44
1	63.53	84.84	97.21	99.15
5	58.30	90.84	98.67	99.57

Na Figura 4, a finalidade é comparar a qualidade do ajuste obtida com a modelagem das despesas médicas entre os métodos de regressão diferencialmente privados e o clássico conforme alteramos o parâmetro de privacidade ε. Para isso, utilizamos a raiz do erro quadrático médio (RMSE) entre os valores reais e preditos, dada por

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \widehat{y}_i)^2},$$

em que $\widehat{y}_i$ são os valores preditos. Observa-se no gráfico que o ajuste não-privado obteve RMSE$=0.097$ e que o método MQO diferencialmente privado apresenta menores magnitudes de erro comparativamente aos demais. Por exemplo para $\varepsilon = 0.01$, o MQO manteve o RMSE próximo de 0.5, enquanto o ajuste de regressão via FM obteve RMSE igual a 68.05 e o DPBA 94.81. Tal fato indica que ambas alternativas da literatura prejudicam fortemente a utilidade do modelo considerando baixos valores de perda de privacidade ε.

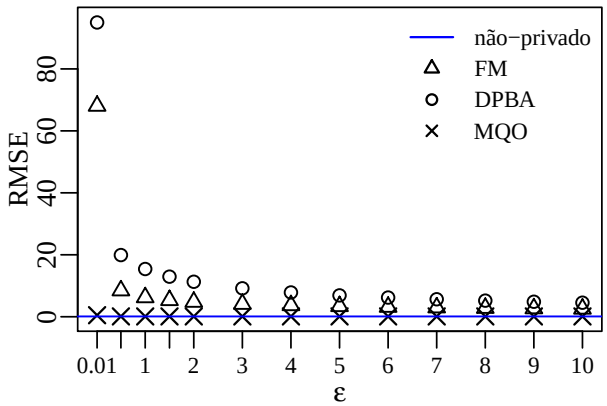


Figura 4 – RMSE $\times$ Privacidade - Despesas médicas ajustadas pela regressão linear.

Selecionamos os modelos em estudo na Tabela 9 para ajustar as despesas médicas de forma privada considerando $\varepsilon = 0.01$. O resultado apresenta os parâmetros estimados, erros padrão, estatística t e seus p-valores para o modelo $y_i = \beta_1 + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_4 x_{i4} + \beta_5 x_{i5} + u_i$. O intercepto é denotado por β_1; os parâmetros β_2 e β_5 são, respectivamente, as covariáveis

idade e fumante, as quais possuem maior significância para descrever a variação das despesas médicas; β_3 representa o índice de massa corporal e β_4 o número de filhos. Foi pertinente considerar apenas variáveis regressoras significativas ao nível de 1% via modelagem clássica (não-privada), no sentido de avaliar o impacto da privacidade causado na significância dos coeficientes estimados. Além disso, consideramos estimadores consistentes ao verificar a presença de heteroscedasticidade no modelo via teste de Breusch-Pagan onde rejeita-se a hipótese nula de homoscedasticidade dado p-valor < 0.001 (BREUSCH; PAGAN, 1979).

Pela Tabela 9, observa-se entre os métodos privados que o MQO e HCs apresentam menores erros, mais proximidade com a modelagem não-privada e preservaram a significância estatística dos parâmetros. Enquanto as alternativas de regressão FM e DPBA obtiveram os erros padrão mais altos para todos parâmetros estimados. Ainda, nota-se que esses dois métodos perturbam fortemente covariáveis limitadas a intervalos menores como é o caso de β_4 (número de filhos) e da variável dummy β_5 (fumante). Da perspectiva dos estimadores consistentes HC0 e HC4, embora os dados apresentam um comportamento heteroscedástico, não houve ganhos expressivos ao considerá-los. Ambos HCs estimam erros padrão muito próximos entre si. Em geral, os resultados assemelham-se aos obtidos utilizando o método tradicional MQO.

5.2.2 Regressão Logística

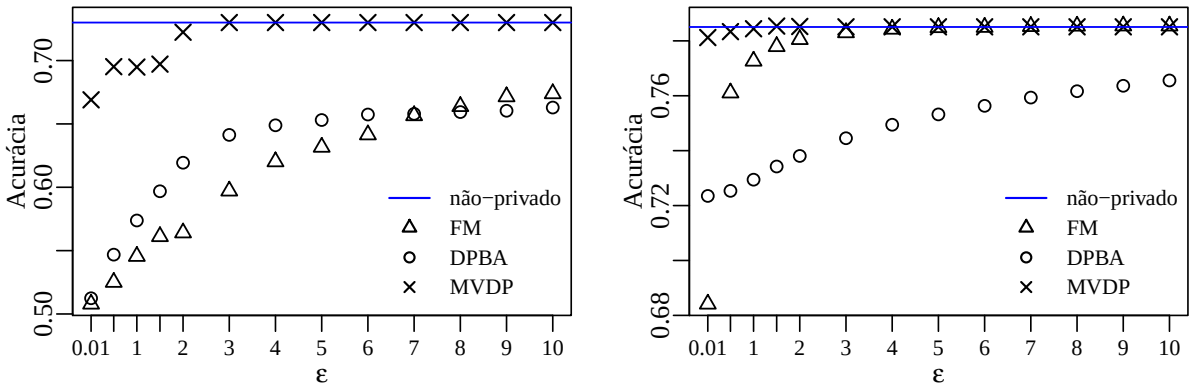
Dados de admissão: O conjunto de dados contendo 400 observações foi obtido da base UCLA - Universidade da Califórnia em Los Angeles[1] e possui uma variável resposta binária que indica, 1 se o aluno foi aprovado ou 0 reprovado em um programa de pós-graduação estadunidense. Desejamos relacionar a admissão com as variáveis independentes do ranqueamento da instituição de proveniência do candidato, bem como suas notas de provas tradicionais nos EUA, GPA (*Grade Point Average*) e GRE (*Graduate Record Examinations*).

Dados da renda adulta: Banco de dados usual na área de preservação de privacidade disponível no repositório UCI Machine Learning (DUA; GRAFF, 2017). Objetiva classificar a renda de adultos com limiar de US$\$ 50,000$ anualmente. Os dados contêm informações demográficas de aproximadamente 47 mil indivíduos como idade, sexo, educação e cargo de trabalho.

A Figura 5 exibe a acurácia de classificação na modelagem dos dados de admissão comparando algoritmos diferencialmente privados a medida que variamos o parâmetro de privacidade ε. A acurácia é medida pela proporção de valores preditos que são equivalentes aos verdadeiros da variável resposta. O modelo privado com perturbação baseada na máxima veros-

[1] Disponível em <https://stats.idre.ucla.edu/>.

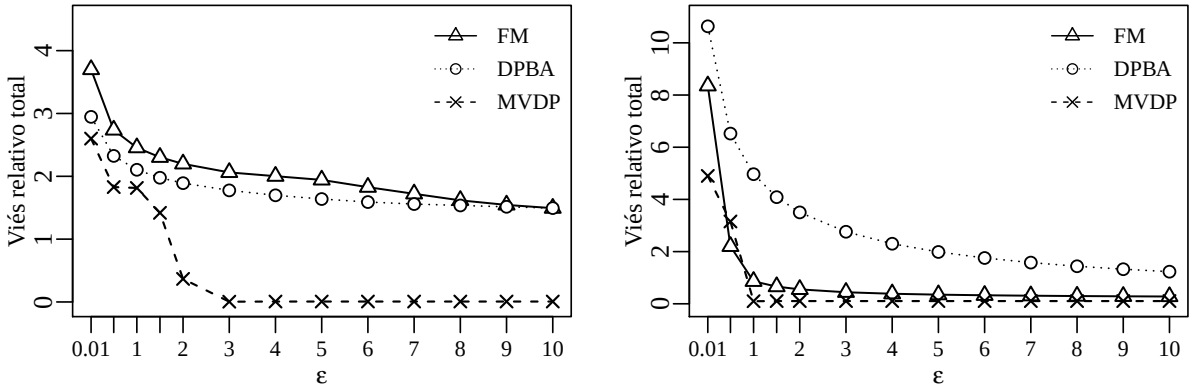
similhança apresenta melhores resultados comparativamente a outras alternativas da literatura em todos níveis de privacidade ε. Isto é, o mecanismo proposto MVDP garante privacidade sem "contaminar" demasiadamente as funções e estimadores da regressão logística.



(a) Dados de admissão **(b) Dados da renda adulta**

Figura 5 – Acurácia $\times$ Privacidade - Dados reais ajustados pela regressão logística.

As abordagens FM e DPBA em ambas Figuras 5a e 5b, comprometem a qualidade do modelo pois necessitam perturbar os dados com maior intensidade para alcançar a proteção das informações pessoais contidas nos dados. Seus desempenhos precários podem ser justificados, pois ao decompor a função objetivo, os algoritmos desconsideram termos polinomiais maiores que de segunda ordem, os quais auxiliam na descrição de características importantes dos dados como assimetria e curtose. No banco de dados da Figura 5b, nota-se que a ponderação do parâmetro ε realizada pelo método DPBA não foi satisfatória, sendo esta mais eficaz nos dados com tamanho menor (5a). Contudo, o algoritmo possibilita alterações ao calibrar a sensibilidade individual, oportunizando o aprimoramento do modelo privado através de novas propostas.



(a) Dados de admissão **(b) Dados da renda adulta**

Figura 6 – Viés relativo total - Dados reais ajustados pela regressão logística.

Na Figura 6, apresentamos os vieses relativos totais para cada método aplicados aos

dados ajustados pela regressão logística. Objetivando a comparação das estimativas privadas com as tradicionais, calculamos o viés relativo de cada coeficiente diferencialmente privado em relação ao seu respectivo obtido pela modelagem clássica (não-privada). Para agregação dos resultados, definimos o viés relativo total como a soma absoluta desses vieses estimados. Em geral, o viés é maior para restrições fortes de privacidade e reduzido conforme a perda de privacidade ε cresce. Verifica-se que o mecanismo de perturbação objetiva baseado na função de log-verossimilhança MVDP apresenta menores distorções comparativamente aos demais mecanismos.

De modo geral, a relação "acurácia $\times$ perda de privacidade" é ratificada nas aplicações dos modelos de regressão diferencialmente privados. Ou seja, para maiores graus de privacidade arbitrado pelo parâmetro ε, mais evidentes são as distorções inferenciais, uma vez que para alcançar tais garantias rigorosas, a perturbação do ruído precisa ser cada vez mais intensa (impondo mais aleatoriedade). Em comparação aos métodos usuais da literatura, destacamos o desempenho superior da proposta MVDP, a qual adiciona incerteza na modelagem de forma mais eficaz para mascarar os dados e fornecer o anonimato aos indivíduos.

Tabela 9 – Modelos de regressão linear para os dados de despesas médicas com $\varepsilon = 0.01$.

	Método	Estimador	Estimativa	Erro padrão	Estatística t	p-valor
β_1	Não-privado	MQO	-0.2110	0.0150	-14.0391	< 0.0001
		HC0		0.0157	-13.4158	< 0.0001
		HC4		0.0158	-13.3415	< 0.0001
	Privado	MQO	-0.0528	0.0161	-3.2764	0.0011
		HC0		0.0173	-3.0599	0.0023
		HC4		0.0174	-3.0387	0.0024
		FM	0.2463	2.2780	0.1081	0.9139
		DPBA	0.3167	3.4289	0.0924	0.9264
β_2	Não-privado	MQO	0.0041	0.0002	21.6746	< 0.0001
		HC0		0.0002	21.8505	< 0.0001
		HC4		0.0002	21.7661	< 0.0001
	Privado	MQO	0.0033	0.0002	16.6195	< 0.0001
		HC0		0.0002	16.3739	< 0.0001
		HC4		0.0002	17.4943	< 0.0001
		FM	0.1906	0.0288	6.6181	< 0.0001
		DPBA	0.2970	0.0433	6.8591	< 0.0001
β_3	Não-privado	MQO	0.0051	0.0004	11.7560	< 0.0001
		HC0		0.0005	10.7694	< 0.0001
		HC4		0.0005	10.6979	< 0.0001
	Privado	MQO	0.0016	0.0005	3.4843	0.0005
		HC0		0.0005	3.1493	0.0016
		HC4		0.0005	3.1553	0.0017
		FM	0.2113	0.0662	3.1918	0.0014
		DPBA	0.3052	0.0997	3.0612	0.0022
β_4	Não-privado	MQO	0.0076	0.0022	3.4364	0.0006
		HC0		0.0021	3.6298	0.0003
		HC4		0.0021	3.6098	0.0003
	Privado	MQO	0.0068	0.0024	2.8892	0.0039
		HC0		0.0022	3.0853	0.0021
		HC4		0.0022	3.0714	0.0022
		FM	0.2462	0.3332	0.7389	0.4600
		DPBA	0.3165	0.5016	0.6309	0.5280
β_5	Não-privado	MQO	0.3801	0.0065	57.9043	< 0.0001
		HC0		0.0091	41.7382	< 0.0001
		HC4		0.0092	41.5350	< 0.0001
	Privado	MQO	0.3121	0.0070	44.3272	< 0.0001
		HC0		0.0106	29.4678	< 0.0001
		HC4		0.0106	29.3017	< 0.0001
		FM	0.2463	0.9945	0.2477	0.8043
		DPBA	0.3166	1.4969	0.2115	0.8325

6 CONSIDERAÇÕES FINAIS

6.1 DISCUSSÕES

O que esperar do estado da arte em privacidade? A escolha ideal de um limiar para o orçamento de privacidade permanece uma questão de pesquisa em aberto que carece mais investigações para aplicabilidade prática. Hsu *et al.* (2014) contribui neste sentido a partir de uma perspectiva econômica. Dwork e Lei (2009) relaciona a seleção do parâmetro de privacidade com uma quantidade de ruído adequado no campo de estatística robusta.

Convenientemente, robustez focada na redução da influência causada por *outliers*, relaciona-se com privacidade quando os dados perturbados não devem afetar expressivamente o resultado ou a distribuição. Ambos conceitos podem ser agregados e alcançados simultaneamente, especialmente, ao utilizar estimadores robustos como ponto de partida para desenvolver procedimentos diferencialmente privados. Por exemplo, ao invés de empregar a função de verossimilhança como função objetivo, aplicar uma função com propriedades robustas como a função de Huber (HUBER, 1981, p. 43).

Outra linha de pesquisa em alta é voltada à maximização da acurácia sem comprometer demasiadamente a privacidade. Muitas técnicas propostas em DP fornecem proteção com grande deterioramento da precisão, portanto há uma necessidade de otimizá-las para melhor desempenho em problemas práticos sob amostras finitas. Trabalhos recentes como Bittau *et al.* (2017) propõem um novo tipo de arquitetura que pode gerar acurácia elevada com baixos níveis de ruído. Essa abordagem chama-se ESA abreviando as seguintes etapas do processo:

***Encoder*:** o codificador coleta os dados e criptografa-os duas vezes;

***Shuffler*:** processo intermediário que remove os identificadores e agrupa dados semelhantes;

***Analyzer*:** descriptografa os dados e calcula as estatísticas de interesse.

Seu funcionamento inicia com o *shuffler*, o qual possui acesso limitado aos identificadores dos indivíduos e não aos dados reais. Ele só pode descriptografar a primeira camada e organizar os dados em grupos. Se houver um número suficiente de usuários nos grupos, estes são passados ao *analyzer* que por sua vez descriptografa a segunda camada e calcula a saída. A novidade está na criptografia separada em duas camadas. A segunda passagem realizada pelo *shuffler* permite que o *analyzer* veja apenas os dados em lotes *(batches)* sem conhecer a origem dos envios. Inserindo aleatoriedade ponderada em locais adequados, tornará o resultado final preservado pela definição de DP.

Há uma perspectiva positiva causada pelo avanço computacional combinado a técnicas de proteção a privacidade. Muitos algoritmos iterativos de aprendizado de máquina sob a restrição de DP estão sendo desenvolvidos. A ideia é ensinar o modelo a capturar informações da distribuição dos dados privados de modo que a dependência individual das informações pessoais seja mínima, isto é, contribuindo para não-exposição da identidade dos indivíduos (GONG *et al.*, 2020). É importante elaborar testes baseados em medidas de exposição ou inferência de associação para certificar que um determinado modelo não memorize informações confidenciais sobre o conjunto de dados. Ainda, demanda-se métodos computacionalmente pouco custosos.

Neste sentido, Papernot *et al.* (2017) propôs o PATE (*Private Aggregation of Teacher Ensembles*), uma estrutura para aprendizagem semi-supervisionada com dados privados mantendo as garantias de DP. A intuição do algoritmo é simples: se diversos classificadores treinados em conjuntos de dados diferentes classificam uma nova entrada da mesma forma, o resultado não revela informações sobre qualquer exemplo de treinamento único e portanto, treinar o modelo com ou sem esse exemplo é indiferente. Recentemente, o PATE foi aprimorado (PAPERNOT *et al.*, 2018) e ampliado para realizar aprendizado profundo via redes neurais (SUN *et al.*, 2019).

O popular algoritmo de aplicabilidade geral *Multiplicative Weights* (LITTLESTONE; WARMUTH, 1994) foi generalizado para preservação de privacidade (citado em 3.6.3). Posteriormente, foi refinado através de uma regra de atualização que pondera positivamente observações mais informativas aliado a um componente ativo de aprendizado baseado no mecanismo exponencial. A proposta obteve ganhos em precisão, tempo de execução e paralelismo para liberação de dados diferencialmente privado (HARDT; LIGETT; MCSHERRY, 2012). Salienta-se que a implementação desses algoritmos não é viável no presente trabalho pois são abordagens voltadas a sistemas *query* e não se adequam a modelos de regressão. As aplicações apresentadas demonstram a evolução e o potencial de privacidade diferencial, ainda que representam uma pequena parcela do que está sendo desenvolvido na área.

Em 2009, a Microsoft lançou o sistema PINQ (*Privacy Integrated Queries*) para análise de dados que preservam a privacidade (de modo diferencialmente privado). A plataforma proporciona maior acessibilidade para realizar aplicações, uma vez que o usuário não necessita experiência ou treinamento especializado. Contudo, o analista é quem decide se o resultado gerado pelo programa é aceitável ou não (MCSHERRY, 2009). A área carece de *softwares* que permitam aplicar as técnicas de DP incluindo composições, análises complexas e combinações de algoritmos. Por exemplo, o usuário especifica sua tarefa e o programa indica os procedimentos mais adequados para executá-la considerando métodos e mecanismos variados. No melhor

cenário, a criação de um sistema para preservação de privacidade com componentes modulares customizáveis e otimização automática promoveria um relevante impacto acadêmico, propício para assegurar o direito à liberdade pessoal.

6.2 CONCLUSÕES E TRABALHOS FUTUROS

Este trabalho apresenta duas abordagens de regressão que satisfazem as garantias de privacidade diferencial através da perturbação objetiva para tornar as informações de um banco de dados não-identificáveis. A estimação dos parâmetros do modelo na regressão linear e logística foi realizada, respectivamente, pelo método de mínimos quadrados ordinários e via máxima verossimilhança; ambos métodos adaptados para incluir a adição de ruído ponderado pela sensibilidade das funções. A avaliação numérica, por meio de simulações de Monte Carlo, evidenciou um crescente viés quanto maior for a restrição de privacidade imposta, assim como maiores distorções nas taxas de rejeição para menores valores do parâmetro de perda de privacidade. Ou seja, o ganho em anonimato é compensado com mais incerteza (em forma de ruído aleatório) no modelo. No caso da regressão linear este resultado é concordante mesmo na presença de heteroscedasticidade. De fato e como era de se esperar, os estimadores consistentes HC0 e HC4 desempenham testes mais poderosos quando comparados ao estimador usual MQO.

Além disso, diferentes concepções de privacidade foram abordadas historicamente sob perspectivas sociais. No âmbito da preservação a privacidade com dados, foi realizada uma ampla revisão literária dos métodos antecedentes e atuais destacando as vantagens que justificam a utilização de DP. Ainda, diversas aplicações industriais e acadêmicas foram relatadas. Por fim, para avaliar empiricamente o desempenho de regressões diferencialmente privadas, a aplicação com dados reais exibe medidas de diagnóstico, as quais confirmam a adequação e superioridade do modelo proposto comparativamente à abordagens usuais da literatura.

Como trabalhos a serem desenvolvidos futuramente, destacamos a inclusão de novos mecanismos que satisfazem as garantias DP como o exponencial e gaussiano para compará-los ao mecanismo de Laplace. Assim como, implementar as perturbações na entrada, saída e no gradiente em modelos de regressão com o propósito comparativo. Avaliar numericamente a aplicação de diferentes métodos para otimização das funções perturbadas, visto que, a inclusão de ruído problematiza a convergência; inclusive considerando derivadas analíticas (quando viável) para analisar como a influência do gradiente afeta as restrições de privacidade DP. Ainda, exerceria grande utilidade a criação de medidas específicas como pontos de ruptura e curvas

de sensibilidade para mensurar analiticamente o impacto de diferentes graus de privacidade no modelo. Para fomentar a reprodutibilidade científica da área, pretende-se estender os modelos diferencialmente privados para a classe dos MLGs, a qual abriga diversas distribuições pertencentes a família exponencial e flexibiliza a escolha da função de ligação.

REFERÊNCIAS

ABADI, M.; CHU, A.; GOODFELLOW, I.; MCMAHAN, H. B.; MIRONOV, I.; TALWAR, K.; ZHANG, L. Deep learning with differential privacy. In: **Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security**. [S.l.: s.n.], 2016. p. 308–318.

ABOWD, J. M. The U.S. Census Bureau adopts differential privacy. In: **Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining**. [S.l.]: Association for Computing Machinery, 2018. p. 2867.

ABOWD, J. M.; VILHUBER, L. How protective are synthetic data? In: SPRINGER, NEW YORK. **Privacy in Statistical Databases**. [S.l.], 2008. p. 239–246.

ACAR, A.; FEREIDOONI, H.; ABERA, T.; SIKDER, A. K.; MIETTINEN, M.; AKSU, H.; CONTI, M.; SADEGHI, A.-R.; ULUAGAC, S. Peek-a-boo: I see your smart home activities, even encrypted! In: **Proceedings of the 13th ACM Conference on Security and Privacy in Wireless and Mobile Networks**. [S.l.]: Association for Computing Machinery, 2020. p. 207–218.

ADAM, N. R.; WORTHMANN, J. C. Security-control methods for statistical databases: a comparative study. **ACM Computing Surveys**, ACM New York, v. 21, n. 4, p. 515–556, 1989.

AGAMBEN, G. **O que é o contemporâneo? E outros ensaios**. [S.l.]: Editora Argos, Chapecó, 2009.

ALEPIS, E.; PATSAKIS, C. Monkey says, monkey does: security and privacy on voice assistants. **IEEE Access**, v. 5, p. 17841–17851, 2017.

ANDREWS, D. W. K. Heteroskedasticity and autocorrelation consistent covariance matrix estimation. **Econometrica**, v. 59, n. 3, p. 817–858, 1991.

APPLE, DIFFERENTIAL PRIVACY TEAM. **Learning with privacy at scale**. 2017. <https://machinelearning.apple.com/research/learning-with-privacy-at-scale>. Acesso em: 22 jul. 2020.

AWAN, J.; SLAVKOVIĆ, A. Structure and sensitivity in differential privacy: comparing k-norm mechanisms. **Journal of the American Statistical Association**, p. 1–20, 2020.

BASSILY, R.; SMITH, A.; THAKURTA, A. Private empirical risk minimization: efficient algorithms and tight error bounds. In: . [S.l.]: IEEE Computer Society, 2014. p. 464–473.

BECK, L. L. A security mechanism for statistical database. **ACM Transactions Database Systems**, v. 5, n. 3, p. 316–338, 1980.

BENTHAM, J. **O Panóptico**. [S.l.]: Autêntica, Belo Horizonte, 2019.

BITTAU, A.; ERLINGSSON, U.; MANIATIS, P.; MIRONOV, I.; RAGHUNATHAN, A.; LIE, D.; RUDOMINER, M.; KODE, U.; TINNES, J.; SEEFELD, B. Prochlo: strong privacy for analytics in the crowd. In: **Proceedings of the 26th Symposium on Operating Systems Principles**. [S.l.: s.n.], 2017. p. 441–459.

BOWEN, C. M.; LIU, F. Comparative study of differentially private data synthesis methods. **Statistical Science**, The Institute of Mathematical Statistics, v. 35, n. 2, p. 280–307, 2020.

BREUSCH, T. S.; PAGAN, A. R. A simple test for heteroscedasticity and random coefficient variation. **Econometrica**, v. 47, n. 5, p. 1287–1294, 1979.

CADWALLADR, C. Google, democracy and the truth about internet search. **The Guardian**, 2016. <https://www.theguardian.com/technology/2016/dec/04/google-democracy-truth-internet-search-facebook>. Acesso em: 11 dez. 2020.

CADWALLADR, C. Revealed: 50 million Facebook profiles harvested for Cambridge Analytica in major data breach. **The Guardian**, 2018. <https://www.theguardian.com/news/2018/mar/17/cambridge-analytica-facebook-influence-us-election>. Acesso em: 03 set. 2020.

CASEY, J. Apple contractors listened to 1,000 siri recordings per shift, says former employee. **Irish Examiner**, 2019. <https://www.irishexaminer.com/news/arid-30945575.html>. Acesso em: 11 nov. 2020.

CHAMIKARA, M.; BERTOK, P.; KHALIL, I.; LIU, D.; CAMTEPE, S. Privacy preserving face recognition utilizing differential privacy. **Computers & Security**, v. 97, 2020.

CHAUDHURI, K.; MONTELEONI, C. Privacy-preserving logistic regression. In: **Advances in neural Information Processing Systems 21**. [S.l.: s.n.], 2009. p. 289–296.

CHAUDHURI, K.; MONTELEONI, C.; SARWATE, A. Differentially private empirical risk minimization. **Journal of Machine Learning Research**, v. 12, n. 3, 2011.

CHAUDHURI, K.; SARWATE, A. D.; SINHA, K. A near-optimal algorithm for differentially-private principal components. **The Journal of Machine Learning Research**, v. 14, n. 1, p. 2905–2943, 2013.

CHEN, K. No place to hide: Edward Snowden, the NSA, and the U.S. surveillance state. **Intelligence and National Security**, v. 32, n. 6, p. 868–871, 2017.

CHEN, M.; MAO, S.; LIU, Y. Big data: a survey. **Mobile Networks and Applications**, v. 19, n. 2, p. 171–209, 2014.

CHIPPERFIELD, J.; GOW, D.; LOONG, B. The australian bureau of statistics and releasing frequency tables via a remote server. **Statistical Journal of the IAOS**, v. 32, n. 1, p. 53–64, 2016.

CHUNG, H.; IORGA, M.; VOAS, J.; LEE, S. Alexa, can I trust you? **Computer**, v. 50, n. 9, p. 100–104, 2017.

CHUNG, H.; LEE, S. Intelligent virtual assistant knows your life. **arXiv preprint arXiv:1803.00466**, 2018.

CORMODE, G.; JHA, S.; KULKARNI, T.; LI, N.; SRIVASTAVA, D.; WANG, T. Privacy at scale: local differential privacy in practice. In: **Proceedings of the 2018 International Conference on Management of Data**. [S.l.: s.n.], 2018. p. 1655–1658.

CRIBARI-NETO, F. Asymptotic inference under heteroskedasticity of unknown form. **Computational Statistics & Data Analysis**, v. 45, n. 2, p. 215–233, 2004.

CRIBARI-NETO, F.; ZARKOS, S. G. Heteroskedasticity-consistent covariance matrix estimation: White's estimator and the bootstrap. **Journal of Statistical Computation and Simulation**, v. 68, n. 4, p. 391–411, 2001.

DAGHER, G. G.; MOHLER, J.; MILOJKOVIC, M.; MARELLA, P. B. Ancile: privacy-preserving framework for access control and interoperability of electronic health records using blockchain technology. **Sustainable Cities and Society**, v. 39, p. 283–297, 2018.

DALENIUS, T. Towards a methodology for statistical disclosure control. **Statistik Tidskrift**, v. 15, n. 429-444, p. 2–1, 1977.

DANDEKAR, A.; BASU, D.; BRESSAN, S. Differential privacy for regularised linear regression. In: **Database and Expert Systems Applications**. [S.l.]: Springer, 2018. p. 483–491.

DAVIDSON, R.; MACKINNON, J. G. **Estimation and Inference in Econometrics**. [S.l.]: Oxford University Press, 1993. (OUP Catalogue).

DELMAZO, C.; VALENTE, J. C. Fake news nas redes sociais online: propagação e reações à desinformação em busca de cliques. **Media & Jornalismo**, v. 18, n. 32, p. 155–169, 2018.

DENNING, D. E.; DENNING, P. J. The tracker: a threat to statistical database security. **ACM Transactions on Database Systems**, v. 4, n. 1, p. 76–96, 1979.

DINUR, I.; NISSIM, K. Revealing information while preserving privacy. In: **Proceedings of the 22nd ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems**. [S.l.: s.n.], 2003. p. 202–210.

DOWELL, J.; CLELAND, J.; FITZPATRICK, S.; MCMANUS, C.; NICHOLSON, S.; OPPÉ, T.; PETTY-SAPHON, K.; KING, O. S.; SMITH, D.; THORNTON, S. *et al.* The UK medical education database what is it? Why and how might you use it? **BMC Medical Education**, v. 18, n. 1, p. 1–8, 2018.

DUA, D.; GRAFF, C. **UCI Machine Learning Repository**. 2017. Disponível em: <http://archive.ics.uci.edu/ml>.

DUCHI, J. C.; JORDAN, M. I.; WAINWRIGHT, M. J. Local privacy and statistical minimax rates. In: **IEEE 54th Annual Symposium on Foundations of Computer Science**. [S.l.: s.n.], 2013. p. 429–438.

DUNCAN, G. T.; LAMBERT, D. Disclosure-limited data dissemination. **Journal of the American statistical association**, v. 81, n. 393, p. 10–18, 1986.

DWORK, C. Differential privacy: a survey of results. In: **Theory and Applications of Models of Computation**. [S.l.: s.n.], 2008. p. 1–19.

DWORK, C.; KENTHAPADI, K.; MCSHERRY, F.; MIRONOV, I.; NAOR, M. Our data, ourselves: privacy via distributed noise generation. In: **Advances in Cryptology - EUROCRYPT**. [S.l.: s.n.], 2006. p. 486–503.

DWORK, C.; LEI, J. Differential privacy and robust statistics. In: **Proceedings of the 41th Annual ACM Symposium on Theory of Computing**. [S.l.: s.n.], 2009. p. 371–380.

DWORK, C.; MCSHERRY, F.; NISSIM, K.; SMITH, A. Calibrating noise to sensitivity in private data analysis. In: **Theory of cryptography conference**. [S.l.: s.n.], 2006. p. 265–284.

DWORK, C.; MCSHERRY, F.; TALWAR, K. The price of privacy and the limits of lp decoding. In: **Proceedings of the 39 annual ACM symposium on Theory of computing**. [S.l.: s.n.], 2007. p. 85–94.

DWORK, C.; ROTH, A. The algorithmic foundations of differential privacy. **Foundations and Trends in Theoretical Computer Science**, v. 9, n. 3-4, p. 211–407, 2014.

DWORK, C.; ROTHBLUM, G. N. Concentrated differential privacy. **arXiv:1603.01887**, 2016.

DWORK, C.; SMITH, A. Differential privacy for statistics: what we know and what we want to learn. **Journal of Privacy and Confidentiality**, v. 1, n. 2, 2010.

EDEMEKONG, P.; ANNAMARAJU, P.; HAYDEL, M. Health Insurance Portability and Accountability Act. **StatPearls**, 2020.

EMAM, K. E.; JONKER, E.; ARBUCKLE, L.; MALIN, B. A systematic review of re-identification attacks on health data. **PLoS One**, v. 6, n. 12, p. 1–12, 2011.

ERLINGSSON, U.; PIHUR, V.; KOROLOVA, A. Rappor: Randomized Aggregatable Privacy-Preserving Ordinal Response. In: **Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security**. [S.l.: s.n.], 2014. p. 1054–1067.

ESTEVA, A.; KUPREL, B.; NOVOA, R. A.; KO, J.; SWETTER, S. M.; BLAU, H. M.; THRUN, S. Dermatologist-level classification of skin cancer with deep neural networks. **Nature**, v. 542, n. 7639, p. 115–118, 2017.

FANG, X.; YU, F.; YANG, G.; QU, Y. Regression analysis with differential privacy preserving. **IEEE Access**, v. 7, p. 129353–129361, 2019.

FELLEGI, I. P. On the question of statistical confidentiality. **Journal of the American Statistical Association**, v. 67, n. 337, p. 7–18, 1972.

FOUCAULT, M. **Discipline and punish: the birth of the prison**. [S.l.]: Vintage, 2012.

FREDRIKSON, M.; JHA, S.; RISTENPART, T. Model inversion attacks that exploit confidence information and basic countermeasures. In: **Proceedings of the 22nd ACM SIGSAC Conference on Computer and Communications Security**. [S.l.: s.n.], 2015. p. 1322–1333.

FREDRIKSON, M.; LANTZ, E.; JHA, S.; LIN, S.; PAGE, D.; RISTENPART, T. Privacy in pharmacogenetics: an end-to-end case study of personalized warfarin dosing. In: **23rd USENIX Security Symposium**. [S.l.: s.n.], 2014. p. 17–32.

FRIEDMAN, A.; SCHUSTER, A. Data mining with differential privacy. In: **Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining**. [S.l.: s.n.], 2010. p. 493–502.

FUKUCHI, K.; TRAN, Q. K.; SAKUMA, J. Differentially private empirical risk minimization with input perturbation. In: **International Conference on Discovery Science**. [S.l.: s.n.], 2017. p. 82–90.

FUNG, B. C. M.; WANG, K.; CHEN, R.; YU, P. S. Privacy-preserving data publishing: a survey of recent developments. **ACM Computing Surveys**, v. 42, n. 4, 2010.

GANTA, S. R.; KASIVISWANATHAN, S. P.; SMITH, A. Composition attacks and auxiliary information in data privacy. In: . [S.l.: s.n.], 2008. p. 265–273.

GARFINKEL, S. L. De-identification of personal information. **National Institute of Standards and Technology**, 2015.

GENG, Q.; VISWANATH, P. The optimal noise-adding mechanism in differential privacy. **IEEE Transactions on Information Theory**, v. 62, n. 2, p. 925–951, 2015.

GERBNER, G.; GROSS, L.; MORGAN, M.; SIGNORIELLI, N. Living with television: the dynamics of the cultivation process. **Perspectives on Media Effects**, p. 17–40, 1986.

GHOSH, A.; ROUGHGARDEN, T.; SUNDARARAJAN, M. Universally utility-maximizing privacy mechanisms. **SIAM Journal on Computing**, v. 41, n. 6, p. 1673–1693, 2012.

GONG, M.; XIE, Y.; PAN, K.; FENG, K.; QIN, A. K. A survey on differentially private machine learning. **IEEE Computational Intelligence Magazine**, v. 15, n. 2, p. 49–64, 2020.

GREENE, W. H. **Econometric Analysis**. 5th. ed. [S.l.]: Pearson Education, 2003.

GREENLEAF, G. Global data privacy laws 2019: 132 national laws & many bills. **157 Privacy Laws & Business International Report**, p. 14–18, 2019.

GROUP, M. M. **Internet World Stats**. 2020. <https://www.internetworldstats.com>. Acesso em: 16 ago. 2020.

HAN, B.-C. **No enxame: perspectivas do digital**. [S.l.]: Editora Vozes, Rio de Janeiro, 2018.

HARDT, M.; LIGETT, K.; MCSHERRY, F. A simple and practical algorithm for differentially private data release. In: **Advances in Neural Information Processing Systems**. [S.l.: s.n.], 2012. p. 2339–2347.

HARDT, M.; ROTHBLUM, G. N. A multiplicative weights mechanism for privacy-preserving data analysis. In: **IEEE 51st Annual Symposium on Foundations of Computer Science**. [S.l.: s.n.], 2010. p. 61–70.

HEATH, S. **Methods and/or systems for an online and/or mobile privacy and/or security encryption technologies used in cloud computing with the combination of data mining and/or encryption of user's personal data and/or location data for marketing of internet posted promotions, social messaging or offers using multiple devices, browsers, operating systems, networks, fiber optic communications, multichannel platforms**. 2018. US Patent 10,129,211.

HEIMANS, J.; TIMMS, H. Understanding "new power". **Harvard Business Review**, v. 92, n. 12, p. 48–56, 2014.

HERE Market Intelligence. **The Privacy Paradox Reloaded: Changes in Consumer Behavior and Attitudes**. 2019. <https://www.here.com/sites/g/files/odxslz166/files/2019-09/the_privacy_paradox.pdf>. Acesso em: 08 out. 2020.

HERE Technologies. **Privacy and Location Data Global Consumer Study**. 2018. https://www.here.com/sites/g/files/odxslz166/files/2019-02/HERE Technologies Privacy and Location Data Global Consumer Study March 2018 - Reviewed.pdf. Acesso em: 08 out. 2020.

HERN, A. New york taxi details can be extracted from anonymised data. **The Guardian**, 2014. <https://www.theguardian.com/technology/2014/jun/27/new-york-taxi-details-anonymised-data-researchers-warn>. Acesso em: 25 out. 2020.

HOLOHAN, N.; BRAGHIN, S.; AONGHUSA, P. M.; LEVACHER, K. Diffprivlib: the IBM differential privacy library. **arXiv:1907.02444**, 2019.

HOOFNAGLE, C.; SOLTANI, A.; GOOD, N.; WAMBACH, D. J. Behavioral advertising: The offer you can't refuse. **Harvard Law and Policy Review**, v. 6, p. 273, 2012.

HSU, J.; GABOARDI, M.; HAEBERLEN, A.; KHANNA, S.; NARAYAN, A.; PIERCE, B. C.; ROTH, A. Differential privacy: an economic method for choosing epsilon. In: **IEEE 27th Computer Security Foundations Symposium**. [S.l.: s.n.], 2014. p. 398–410.

HUBER, P. J. **Robust Statistics**. [S.l.]: United States of America, John Wiley & Sons, 1981. Wiley Series in Probability and Statistics.

ITU. **Measuring digital development: facts and figures**. 2019. International Telecommunication Union. <https://itu.foleon.com/itu/measuring-digital-development>. Acesso em: 30 jul. 2020.

JAIN, P.; GYANCHANDANI, M.; KHARE, N. Big data privacy: a technological perspective and review. **Journal of Big Data**, v. 3, n. 1, p. 1–25, 2016.

JAIN, P.; GYANCHANDANI, M.; KHARE, N. Differential privacy: its technological prescriptive using big data. **Journal of Big Data**, v. 5, n. 1, p. 1–15, 2018.

JIANG, X.; JI, Z.; WANG, S.; MOHAMMED, N.; CHENG, S.; OHNO-MACHADO, L. Differential-private data publishing through component analysis. v. 6, n. 1, p. 19–34, 2013.

JOHNSON, N.; NEAR, J. P.; SONG, D. Towards practical differential privacy for SQL queries. **Proceedings of the VLDB Endowment**, v. 11, n. 5, p. 526–539, 2018.

JOSHI, A. P.; HAN, M.; WANG, Y. A survey on security and privacy issues of blockchain technology. **Mathematical Foundations of Computing**, v. 1, n. 2, p. 121, 2018.

KANG, Y.; LIU, Y.; NIU, B.; TONG, X.; ZHANG, L.; WANG, W. Input perturbation: a new paradigm between central and local differential privacy. **arXiv:2002.08570**, 2020.

KELION, L. Samsung's smart tvs fail to encrypt voice commands. **BBC News**, 2015. <https://www.bbc.com/news/technology-31523497>. Acesso em: 21 nov. 2020.

KIFER, D.; SMITH, A.; THAKURTA, A. Private convex empirical risk minimization and high-dimensional regression. In: **25th Annual Conference on Learning Theory**. [S.l.: s.n.], 2012. v. 23, p. 25.1–25.40.

KINGMA, D. P.; WELLING, M. Auto-encoding variational bayes. **arXiv:1312.6114**, 2013.

KOHANE, I. S. Using electronic health records to drive discovery in disease genomics. **Nature Reviews Genetics**, v. 12, n. 6, p. 417–428, 2011.

LANTZ, B. **Machine Learning with R**. [S.l.]: Packt Publishing, 2019.

LEE, J.; CLIFTON, C. How much is enough? choosing ε for differential privacy. In: **International Conference on Information Security**. [S.l.: s.n.], 2011. p. 325–340.

LENSVELT-MULDERS, G.; HOX, J.; HEIJDEN, P. V. D. How to improve the efficiency of randomised response designs. **Quality and Quantity**, v. 39, n. 3, p. 253–265, 2005.

LI, C.; MIKLAU, G.; HAY, M.; MCGREGOR, A.; RASTOGI, V. The matrix mechanism: optimizing linear counting queries under differential privacy. **The VLDB Journal**, v. 24, n. 6, p. 757–781, 2015.

LI, Y.; DAI, W.; MING, Z.; QIU, M. Privacy protection for preventing data over-collection in smart city. **IEEE Transactions on Computers**, v. 65, n. 5, p. 1339–1350, 2015.

LITTLESTONE, N.; WARMUTH, M. The weighted majority algorithm. **Information and Computation**, v. 108, n. 2, p. 212–261, 1994.

LIU, F. Generalized gaussian mechanism for differential privacy. **IEEE Transactions on Knowledge and Data Engineering**, v. 31, n. 4, p. 747–756, 2019.

LO'AI, A. T.; MEHMOOD, R.; BENKHLIFA, E.; SONG, H. Mobile cloud computing model and big data analysis for healthcare applications. **IEEE Access**, v. 4, p. 6171–6180, 2016.

LONG, J. S.; ERVIN, L. H. Using heteroscedasticity consistent standard errors in the linear regression model. **The American Statistician**, v. 54, n. 3, p. 217–224, 2000.

LYU, L.; NANDAKUMAR, K.; RUBINSTEIN, B.; JIN, J.; BEDO, J.; PALANISWAMI, M. PPFA: Privacy Preserving Fog-enabled Aggregation in smart grid. **IEEE Transactions on Industrial Informatics**, v. 14, n. 8, p. 3733–3744, 2018.

MACHADO, J. d. M. S. A expansão do conceito de privacidade e a evolução na tecnologia de informação com o surgimento dos bancos de dados. **Revista da AJURIS**, v. 41, n. 134, 2014.

MACHANAVAJJHALA, A.; KIFER, D.; ABOWD, J.; GEHRKE, J.; VILHUBER, L. Privacy: theory meets practice on the map. In: **IEEE 24th International Conference on Data Engineering**. [S.l.: s.n.], 2008. p. 277–286.

MACHANAVAJJHALA, A.; KIFER, D.; GEHRKE, J.; VENKITASUBRAMANIAM, M. L-diversity: privacy beyond k-anonymity. **ACM Transactions on Knowledge Discovery from Data**, v. 1, n. 1, p. 3–es, 2007.

MACKINNON, J. G.; WHITE, H. Some heteroskedasticity-consistent covariance matrix estimators with improved finite sample properties. **Journal of Econometrics**, v. 29, n. 3, p. 305–325, 1985.

MCSHERRY, F. Privacy integrated queries: an extensible platform for privacy-preserving data analysis. In: **Proceedings of the 2009 ACM SIGMOD International Conference on Management of Data**. [S.l.: s.n.], 2009. p. 19–30.

MCSHERRY, F.; MIRONOV, I. Differentially private recommender systems: building privacy into the netflix prize contenders. In: **Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining**. [S.l.]: Association for Computing Machinery, 2009. p. 627–636.

MCSHERRY, F.; TALWAR, K. Mechanism design via differential privacy. In: **48th Annual IEEE Symposium on Foundations of Computer Science**. [S.l.: s.n.], 2007. p. 94–103.

MEIRELLES, FGV. **31ª Pesquisa Anual do Uso de Tecnologia da Informação nas Empresas**. 2020. <http://www.fgv.br/cia/pesquisa>. Fundação Getulio Vargas: Centro de TI Aplicada da EAESP.

NARAYANAN, A.; SHMATIKOV, V. Robust de-anonymization of large sparse datasets. In: IEEE. **2008 IEEE Symposium on Security and Privacy (sp 2008)**. [S.l.], 2008. p. 111–125.

NEEL, S.; ROTH, A.; VIETRI, G.; WU, Z. S. Oracle efficient private non-convex optimization. In: **Proceedings of the 37th International Conference on Machine Learning**. [S.l.: s.n.], 2019.

NEWEY, W. K.; WEST, K. D. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. **Econometrica**, v. 55, n. 3, p. 703–708, 1987.

NEWEY, W. K.; WEST, K. D. Automatic lag selection in covariance matrix estimation. **The Review of Economic Studies**, v. 61, n. 4, p. 631–653, 1994.

NISSIM, K.; RASKHODNIKOVA, S.; SMITH, A. Smooth sensitivity and sampling in private data analysis. In: **Proceedings of the 39th ACM Symposium on Theory of Computing**. [S.l.: s.n.], 2007. p. 75–84.

NOCEDAL, J.; WRIGHT, S. **Numerical Optimization**. [S.l.]: Springer, New York, 1999.

NORBERG, P. A.; HORNE, D. R.; HORNE, D. A. The privacy paradox: personal information disclosure intentions versus behaviors. **Journal of Consumer Affairs**, v. 41, n. 1, p. 100–126, 2007.

OBAR, J. A.; OELDORF-HIRSCH, A. The biggest lie on the internet: ignoring the privacy policies and terms of service policies of social networking services. **Information, Communication & Society**, v. 23, n. 1, p. 128–147, 2020.

OHM, P. Broken promises of privacy: responding to the surprising failure of anonymization. **UCLA Law Review**, v. 57, p. 1701–1777, 2009.

PAPERNOT, N.; ABADI, M.; ERLINGSSON, U.; GOODFELLOW, I.; TALWAR, K. Semi-supervised knowledge transfer for deep learning from private training data. In: **Proceedings of the 5th International Conference on Learning Representations**. [S.l.: s.n.], 2017.

PAPERNOT, N.; SONG, S.; MIRONOV, I.; RAGHUNATHAN, A.; TALWAR, K.; ERLINGSSON, Ú. Scalable private learning with PATE. In: **Proceedings of the 6th International Conference on Learning Representations**. [S.l.: s.n.], 2018.

PATHAK, M. A.; RAJ, B. Large margin gaussian mixture models with differential privacy. **IEEE Transactions on dependable and secure computing**, v. 9, n. 4, p. 463–469, 2012.

PAULA, G. A. **Modelos de regressão com apoio computacional**. 2013. <https://www.ime.usp.br/~giapaula/texto_2013.pdf>. Acesso em: 13 ago. 2020.

PAWITAN, Y. **In All Likelihood: Statistical Modelling and Inference Using Likelihood**. [S.l.]: Oxford University Press, 2001.

PERSE, E. M.; LAMBE, J. **Media effects and society**. [S.l.]: Routledge, Taylor & Francis Group, New York, 2016.

PFEIFLE, A. Alexa, what should we do about privacy? protecting privacy for users of voice-activated devices. **Washington Law Review**, v. 93, p. 421, 2018.

PHAN, N.; WU, X.; HU, H.; DOU, D. Adaptive laplace mechanism: differential privacy preservation in deep learning. In: **2017 IEEE International Conference on Data Mining**. [S.l.: s.n.], 2017. p. 385–394.

Pierre-Simon Laplace. Mémoire sur la probabilité des causes par les événements. **Mémoires de l'Academie Royale des Sciences**, v. 6, p. 621–656, 1774.

PRICE, N.; COHEN, G. Privacy in the age of medical big data. **Nature Medicine**, v. 25, n. 1, p. 37–43, 2019.

R Core Team. **R: A Language and Environment for Statistical Computing**. Vienna, Austria, 2020. Disponível em: <https://www.R-project.org/>.

RAGHUNATHAN, T. E.; REITER, J. P.; RUBIN, D. B. Multiple imputation for statistical disclosure limitation. **Journal of Official Statistics**, v. 19, n. 1, p. 1, 2003.

REN, J.; DUBOIS, D. J.; CHOFFNES, D.; MANDALARI, A. M.; KOLCUN, R.; HADDADI, H. Information exposure from consumer IoT devices: a multidimensional, network-informed measurement approach. In: **Proceedings of the Internet Measurement Conference**. [S.l.: s.n.], 2019. p. 267–279.

REUTER, M. S.; WALKER, S.; THIRUVAHINDRAPURAM, B.; WHITNEY, J.; COHN, I.; SONDHEIMER, N.; YUEN, R. K.; TROST, B.; PATON, T. A.; PEREIRA, S. L. *et al.* The personal genome project canada: findings from whole genome sequences of the inaugural 56 participants. **Canadian Medical Association Journal**, v. 190, n. 5, p. E126–E136, 2018.

REVOREDO, T. **Blockchain: tudo o que você precisa saber**. [S.l.]: Amazon US, 2019.

Risk Based Security. **Data Breach Report 2019 MidYear QuickView**. 2019. <https://pages.riskbasedsecurity.com/2019-midyear-data-breach-quickview-report>. Acesso em: 01 set. 2020. Cyber Risk Analytics.

Risk Based Security. **Data Breach QuickView 2020 Mid Year Report**. 2020. <https://pages.riskbasedsecurity.com/en/2020-mid-year-data-breach-quickview-report>. Acesso em: 01 set. 2020. Cyber Risk Analytics.

ROCHER, L.; HENDRICKX, J. M.; MONTJOYE, Y.-A. D. Estimating the success of re-identifications in incomplete datasets using generative models. **Nature Communications**, v. 10, 2019.

RODOTÀ, S. A vida na sociedade da vigilância: a privacidade hoje. In: . [S.l.]: Renovar, 2015.

ROTH, A.; ROUGHGARDEN, T. Interactive privacy via the median mechanism. In: **Proceedings of the 42nd ACM Symposium on Theory of Computing**. [S.l.: s.n.], 2010. p. 765–774.

RUBIN, D. B. Discussion statistical disclosure limitation. **Journal of Official Statistics**, v. 9, n. 2, p. 461–468, 1993.

SANKAR, L.; RAJAGOPALAN, S. R.; POOR, H. V. Utility-privacy tradeoffs in databases: a information-theoretic approach. **IEEE Transactions on Information Forensics and Security**, v. 8, n. 6, p. 838–852, 2013.

SARATHY, R.; MURALIDHAR, K. Evaluating laplace noise addition to satisfy differential privacy for numeric data. **Transactions on Data Privacy**, v. 4, n. 1, p. 1–17, 2011.

SARWATE, A. D.; CHAUDHURI, K. Signal processing and machine learning with differential privacy: algorithms and challenges for continuous data. **IEEE Signal Processing Magazine**, v. 30, n. 5, p. 86–94, 2013.

SCHÖNBERGER, V.; CUKIER, K. **Big data: a revolution that will transform how we live, work, and think**. [S.l.]: Houghton Mifflin Harcourt, Boston, 2013.

SHOKRI, R.; STRONATI, M.; SONG, C.; SHMATIKOV, V. Membership inference attacks against machine learning models. In: **IEEE Symposium on Security and Privacy**. [S.l.: s.n.], 2017. p. 3–18.

SNOKE, J.; RAAB, G.; NOWOK, B.; DIBBEN, C.; SLAVKOVIC, A. General and specific utility measures for synthetic data. **Journal of the Royal Statistical Society. Series A: Statistics in Society**, v. 181, n. 3, p. 663–688, 2018.

SOLOVE, D. J. Understanding privacy. Harvard University Press, 2008.

SONG, S.; CHAUDHURI, K.; SARWATE, A. D. Stochastic gradient descent with differentially private updates. In: **2013 IEEE Global Conference on Signal and Information Processing**. [S.l.: s.n.], 2013. p. 245–248.

SUN, L.; ZHOU, Y.; WANG, J.; LI, J.; SOCHAR, R.; YU, P. S.; XIONG, C. Private deep learning with teacher ensembles. **arXiv:1906.02303**, 2019.

SWEENEY, L. k-anonymity: a model for protecting privacy. **International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems**, v. 10, n. 5, p. 557–570, 2002.

TADDICKEN, M. The privacy paradox in the social web: the impact of privacy concerns, individual characteristics, and the perceived social relevance on different forms of self-disclosure. **Journal of Computer-Mediated Communication**, v. 19, n. 2, p. 248–273, 2014.

TRAUB, J. F.; YEMINI, Y.; WOŹNIAKOWSKI, H. The statistical security of a statistical database. **ACM Transactions on Database Systems**, v. 9, n. 4, p. 672–679, 1984.

UNITED NATIONS. **The Sustainable Development Goals Report**. 2020. <https://unstats.un.org/sdgs/report/2020>. United Nations Statistics Division.

United Nations General Assembly. Universal declaration of human rights. **Resolution 217 A, art.12**, 1948.

WAHLBERG, A.; SJOBERG, L. Risk perception and the media. **Journal of Risk Research**, v. 3, n. 1, p. 31–50, 2000.

WANG, B.; LI, X. Big data, platform economy and market competition: a preliminary construction of plan-oriented market economy system in the information era. **World Review of Political Economy**, v. 8, n. 2, p. 138–161, 2017.

WANG, Y.; SI, C.; WU, X. Regression model fitting under differential privacy and model inversion attack. In: **International Joint Conferences on Artificial Intelligence**. [S.l.: s.n.], 2015. p. 1003–1009.

WARNER, S. L. Randomized response: a survey technique for eliminating evasive answer bias. **Journal of the American Statistical Association**, v. 60, n. 309, p. 63–69, 1965.

WARREN, S. D.; BRANDEIS, L. D. The right to privacy. **Harvard Law Review**, The Harvard Law Review Association, v. 4, n. 5, p. 193–220, 1890.

WASSERMAN, L.; ZHOU, S. A statistical framework for differential privacy. **Journal of the American Statistical Association**, v. 105, n. 489, p. 375–389, 2010.

WESTIN, A. **Privacy and Freedom**. [S.l.]: Editora Atheneum, New York, 1967.

WHITE, H. A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. **Econometrica**, v. 48, n. 4, p. 817–838, 1980.

WILLENBORG, L.; WAAL, T. D. **Elements of Statistical Disclosure Control**. [S.l.]: Springer Science & Business Media, 2012. v. 155.

WINDER, D. **Confirmed: 2 Billion Records Exposed In Massive Smart Home Device Breach**. 2019. <https://www.forbes.com/sites/daveywinder/2019/07/02/confirmed-2-billion-records-exposed-in-massive-smart-home-device-breach>. Acesso em: 27 nov. 2020.

WOODWARD, B. The computer-based patient record and confidentiality. **Advancing the Consumer Interest**, Wiley, v. 8, n. 1, p. 18–22, 1996.

WU, X.; LI, F.; KUMAR, A.; CHAUDHURI, K.; JHA, S.; NAUGHTON, J. Bolt-on differential privacy for scalable stochastic gradient descent-based analytics. In: **Proceedings of the 2017 ACM International Conference on Management of Data**. [S.l.: s.n.], 2017. p. 1307–1322.

YU, F.; FIENBERG, S. E.; SLAVKOVIC, A. B.; UHLER, C. Scalable privacy-preserving data sharing methodology for genome-wide association studies. **Journal of Biomedical Informatics**, v. 50, p. 133 – 141, 2014.

YUE, X.; WANG, H.; JIN, D.; LI, M.; JIANG, W. Healthcare data gateways: found healthcare intelligence on blockchain with novel privacy risk control. **Journal of Medical Systems**, v. 40, n. 10, p. 218, 2016.

ZHANG, J.; ZHANG, Z.; XIAO, X.; YANG, Y.; WINSLETT, M. Functional mechanism: regression analysis under differential privacy. **Proceedings of the VLDB Endowment**, v. 5, n. 11, 2012.

ZHANG, Z.; RUBINSTEIN, B.; DIMITRAKAKIS, C. On the differential privacy of bayesian inference. **arXiv:1512.06992**, 2015.

ZIGOMITROS, A.; CASINO, F.; SOLANAS, A.; PATSAKIS, C. A survey on privacy properties for data publishing of relational data. **IEEE Access**, v. 8, p. 51071–51099, 2020.

ZUBOFF, S. Big other: surveillance capitalism and the prospects of an information civilization. **Journal of Information Technology**, v. 30, n. 1, p. 75–89, 2015.

ZYSKIND, G.; NATHAN, O.; PENTLAND, A. Decentralizing privacy: using blockchain to protect personal data. In: **2015 IEEE Security and Privacy Workshops**. [S.l.: s.n.], 2015. p. 180–184.