Università degli Studi di Padova

Dipartimento di Scienze Statistiche

Corso di Laurea Magistrale in
Scienze Statistiche

The protection of personal data: a theoretical approach via differential privacy

Relatore:
Prof. Marco Formentin
Dipartimento di Matematica
"Tullio Levi-Civita"

Laureando:
Sofia Ferrari
Matricola: 2026885

Anno Accademico 2021/2022

Contents

Introduction

The increasing use of Internet in daily life, such as for work, socialising, health care and other essential services, has led to an ever larger release of data by users.
For this reason and after large companies or healthcare institutions have been victims of data breaches and compromises, the issue of protecting the individual has become crucial.
This is why it is necessary to introduce a privacy system that can protect the individual in the situation where his data is in the hands of a person who wants to know more about him without his consent.
In this thesis we deal with a type of protection that caught on in the early 2000s and that can be a good solution for providing security: the *differential privacy*.
We wanted to present and explore this topic first of all because of the importance of the topic nowadays and to explore the mechanisms and limitations that characterise it. The aim of the thesis is in fact to develop differential privacy from a theoretical point of view using definitions and theorems of a mathematical statistical nature.
Subsequently, through practical applications, the differentiated privacy of some mechanisms was demonstrated, using the Rstudio software.
In the first chapter of this thesis, the problem of privacy on the Internet is introduced, accompanied by related real-life examples of violations in recent years in various fields.
Based on this primary need, the European Union introduced a regulation, the GDPR, to guarantee the confidentiality of an individual's information as a right of every person.
In order to help with these problems, the common solution is differential privacy.
This process is introduced by three desirdata and then some examples of application

by organisations in recent years are presented.

The output returned by a differentiated privacy process prevents from identifying whether or not an individual participated in the study. It ensures that the addition or removal of an individual in a dataset does not change the differential privacy guarantee given.

The main characteristic of the output obtained with a differentiated privacy process is the addition of noise to the initial data, resulting an approximation of the original output. The privacy loss parameter, ϵ, smaller it is, more the data are protected but less accurate, and vice versa.

The aim is to find the right compromise between data security and adherence to reality in order to use them.

The fundamental role of ϵ and the form it can take is dealt with in the second chapter, starting with the formal definition of differential privacy and continuing with the mechanisms behind it.

Examples of this are, randomised responses and the Laplace mechanism, which plays a fundamental role and which we will also find in later chapters.

The last mechanism assumes that the error follows a Laplace distribution, which can be applied to various case studies.

It will be seen how an individual's information can be viewed under a Bayesian approach, taking what is available to an individual as an a priori distribution.

Often, however, differential privacy applies more error than it needs to in reality, which is why assumptions are relaxed and the definition of privacy is weakened.

In fact, two examples that follow this line, are mentioned: (ϵ, δ)-differential privacy and (ϵ, α)-Rényi differential privacy.

We will focus in particular on the second type of privacy due to its characteristics and properties.

Some examples are: resistance to additional information, group privacy and the important property of the composition that guarantee the privacy to users who have participated in several studies.

Again, privacy is applied to the examples seen above with the addition of the Gaussian

mechanism, assuming that the error follows a Normal distribution.

The final part of the thesis was dedicated to collecting the concepts discussed above and applying them to datasets by proving their differential privacy using special algorithms: Laplace, Gaussian and Bernstein mechanisms.

We used the 'diffpriv' package for the analysis, reporting results and graphs.

Chapter 1

Introduction to the Differential Privacy

1.1 The two sides of information

The recent evolution of technologies and the consequent increase of digital users, have generated an exponential increase of information during the last years.
The circulation of data is caused by simple actions that anyone does every day like: searching on the Web, allowing to privacy requests, watching streaming movies, using of maps, interacting on social networks and many other behaviors of daily life.
The collection of this data fills databases that, using specific algorithms, are used for multiple purposes: an example is the knowledge of consumer preferences and user's habits.
The explosion of data quantity, the increase of the complexity and the speed of transmission have led to the incurring of high costs related to the necessity to store more data. This led to the development of new methods of collecting them, defining a new category of data, the *Big Data*.
The *Big Data* have generated progress in many areas: thinking, for example, of the geolocation of people, the support in the medical field or, again, the prediction of extraordinary events.

1.1.1 The role of privacy

The constant increasing of digital data has generated numerous advantages, but by the other hand has strongly contributed to the diffusion of sensitive data.
This phenomenon, which involved all the World, has compromised users' rights, causing cases of privacy violation .
The unstoppable development of this phenomenon has favored the diffusion of information about the identity of people, such as their name, place and date of birth, and also their personal information, like marital status, occupation, religion and ethnicity.

1.1.2 The legislation of EU: GDPR

The increasing risk of violation of people identity and their information, has caused the introduction of a general regulation for data protection, in 24 May 2016: the General Data Protection Regulation or GDPR[1].
This regulation has been applied in the European Union only two years later, for the security of its citizens.
The main point of this regulation is that the protection of personal data becomes a right for the safety of the population.
Other important points that are introduced for the first time in the European Union, through the GDPR. These issues are: the limitation of the amount of data used for the study involved, the time of data use must not be higher than necessary, the purpose of the data used and their security.
What emerges is that the information available are used for what is really necessary, guaranteeing users about their safety and protecting them from "external" attacks.
If on one side you want to guarantee the people's privacy, on the other the information are useful to produce statistical analysis results, or as mentioned above, even for companies or for other business purposes.

[1]Official Journal of the European Union, 27 April 2016: *"The protection of individuals with regard to the processing of personal data is a fundamental right. Article 8, paragraph 1, of the Charter of Fundamental Rights of the European Union ("Charter") and Article 16 (1) of the Treaty on the Functioning of the European Union ("TFEU") establish that everyone has the right to the protection of personal data concerning them. "*

In order to collect information about people's answer, there must still be the consent of user, often not conscious of the use that "external" individuals can make of it.
What often happens is that if these responses are combined with global domain data collections, a single individual's identity and sensitive data can be violated in any case. This aspect will be explored in the following paragraphs and the term "participant" will be used to indicate the person of interest who releases personal data, while "adversary" for the one who uses the data without a specific purpose.

1.2 Examples of privacy weaknesses

We present below some examples where users' privacy, without difficulty, can be compromised in different areas, starting from a simple online movie review to data regarding an users' health.

1.2.1 Netflix's challenge

The first example that is taken in this analysis is the case of Netflix which is presented in the article [4].
In 2006, the American company that offers a streaming service for TV programmes and films online, after publishing a ranking of the platform's most streamed movies by users, challenged experts and lovers of recommendation system development.
Arvind Narayanan and Vitaly Shmatikov, two researchers from the University of Texas took up the challenge. They proved that with this data it is possible to trace individual subjects despite the fact that Netflix, in order to protect users, replaced names through randomization of numbers and anonymization of personal data collected.
An "adversary," could trace the subscriber through additional information such as: last movies watched, genre preferences and dates they watched TV series or movies.
The technique used to remove anonymity of the study's participants was the introduction of time markers attached to the initially available data, relating to the viewing history of movies and then be able to find the original dataset.

The additional information, used to trace back to the originals data, was sourced from *Internet Movies Database* (IMDb), a public domain online databank where statistics and information on movies, TV series, video games, and many others can be found.
From this database, users who gave reviews by making their names explicit, were randomly chosen to be later compared with the data made public by Netflix.
Specifically, they deleted the 100 most viewed movies and they analyzed the other in order to make more easier to find the attitude taken by the individual subjects and their identity.
At this point, with only 8 movie reviews with the attached dates on which it was reviewed we can identify users. The 99% of the users can be uniquely identified, with a margin of error of 14 days, while 68% of the dataset can be identified with only two movie reviews and dates (with a margin of error of 3 days), as observed in Figure 1.1.
It was seen that easy-to-find items such as a public, small database were used for this analysis leading to a high risk of replication of the process in many other cases.
The solution for this analysis would seem to modify the temporal markers or randomize the data by removing some of it from the set. The two researchers have shown how these aspects are useful only to complicate the identification of subjects and not to protect them totally.
In fact, their algorithm, again, is shown to be workable and robust.
If only the available database released, as in this case, by Netflix were used the informations that could be derived about an individual is almost nil, but if auxiliary indications are available, these can create problems for security on a user's data.
A suitable technique for solving this problem is an innovative solution called, *differential privacy.* This privacy tells us that even in the presence of correlation between a user in the dataset and a given external database, the "adversary" is not able to know any additional information beyond what it already has.

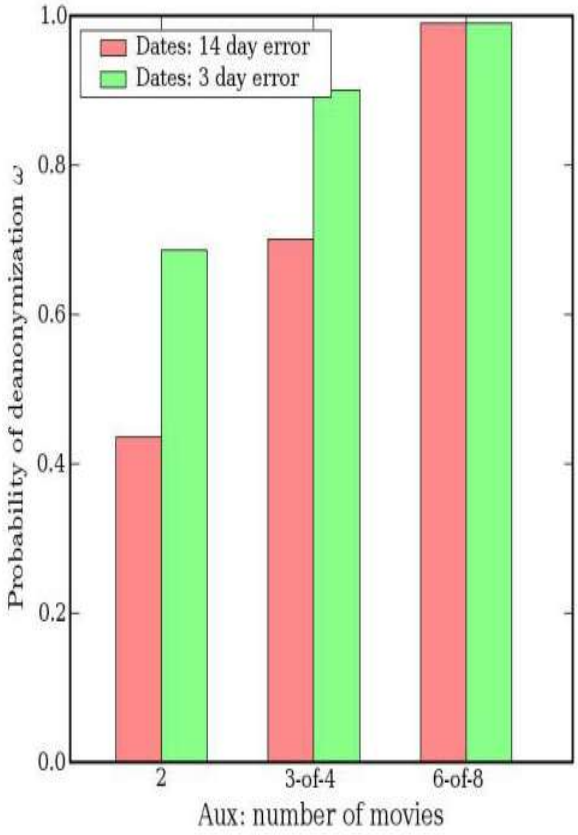


Figure 1.1: *Frequency histogram of the probability of identifying the participation in the study, given the exact film scores and approximate dates taken from [10].*

1.2.2 The difficult protection of health data

1.2.3 Nils Homer and genomics

The development of technologies, innovations, and continued studies have shown a crucial role in the advancement of the health care field, which is presented in the article [11].

Contributing to this growth is the role of data and information that has been collected

over time and continues to be stored for use in new studies or treatments.

The focus here is on a branch of biology that studies the genome of organisms: genomics.

The U.S. National Library of Medicine's, "Database of Genotypes and Phenotypes" (dbGaP) is one of the data collections that has become increasingly relevant within this discipline over the years.

The National Human Genome Research Institute, in fact, compared the studies done on the genome and the results obtained carried out in three different years: 2005, 2010 and 2013. For each year counting less than 50 studies in 2005, to 900 in 2010 and finally 2000 studies in 2003, expanding more and more the data collections available.

The constant increase in information, has also elevated concerns and the possibility of breaching an individual's sensitive data, speaking in this case of his or her health.

The bio-informatician, Nils Homer, published a paper in 2008 where he exposes this problem, emphasizing the fact that with the genome data of an unknown subject, it was shown that it could be traced back to the identification of people's identity.

Seeing the danger, the Institute of Health (NIH) and other health institutes instantaneously removed genome information from the public domain. They later introduced greater care in the granting of data collected in the database through stricter and more controlled regulations, which are still in effect today.

The question has been raised on how to be able to prevent attacks from "adversaries" on genomic data, considering the low numerosity in data collections, the high correlation between data, and the high amount of different databases concerning the health field.

Simple privacy based on anonymity is not enough, as researchers Narayanan and Shmatikov have shown. For example, in the case of Netflix, because the cross-referencing of different datasets is sufficient to trace the individual, although much data has been masked for public data protection.

One solution, again, is through *differential privacy*, which provides the "participant" in the study with greater protection.

Like the previous case, due to the underlying logic of this protection, an "adversary" will not be able to obtain additional information about the "participant" than what he

or she already possessed previously.

Complete privacy cannot be achieved, but a compromise is sought between good coverage of this and use of the data for the purpose of study or statistical analysis.

As an alternative to *differential privacy*, in this case, one element that can be introduced in favor of data protection is a privacy optimization parameter based on the utility and nature of the data.

1.2.4 Evolution of privacy in Australia

We now study the case of a geographical area that has demonstrated poor protection of its residents' personal information during the years: Australia. The South Australia Health, has made public a chart concerning hospital data realtivized respiratory and intestinal infections between 2005 and 2018, whose subjects "participants" in the study are children.

What was published, was linked to the original data by adding the subjects' first name, last name and date of birth to the chart.

The non-immediate detection and intervention of the data violation on the network, gave 300 people the opportunity to view this information, compromising the privacy of the subjects present, particularly of underage "participants".

It is intended to present a case in which the Australian population again encountered the personal data breach, with the difference that in the next example, the data was shared after applying measures to ensure privacy.

Australia's Department of Health, published in 2016 information regarding the medical billing records of a sample of the Australian population, in this case 2,985,511 people. The information was found from the website "data.gov.au" and concern the state of health of a portion of the Australian population.

The datasets were downloaded more than a thousand times, unaware of the little protection afforded to users participating in the analysis.

Only a couple of years later it was discovered, thanks to the University of Melbourne, that the initial data could be traced simply by finding a weakness in the algorithm

used to protect the underlying data, confirming the fact that the tools available were not sufficient.

The original data were never published but were only used as evidence of weakness in the underlying database protection system.

A solution related to *differential privacy* is used, again like the previous examples.

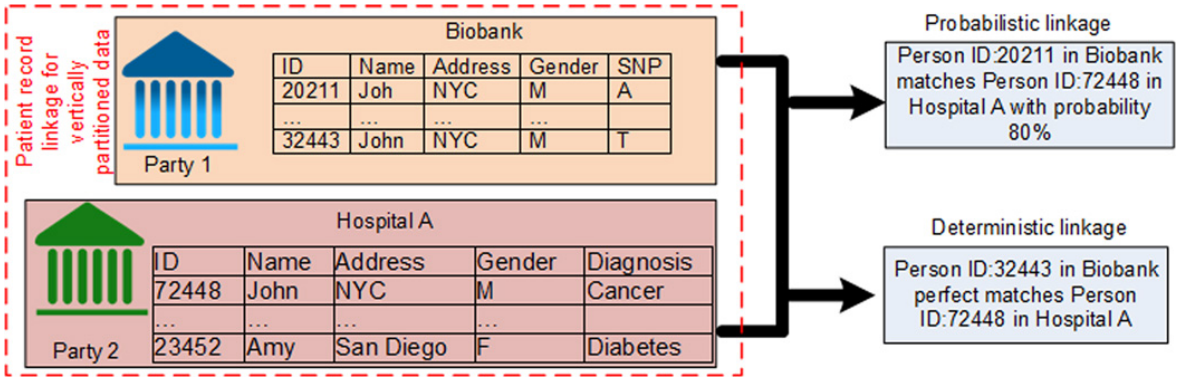


Figure 1.2: *Identification of an individual through a database and additional information taken from [12]*

1.3 A common solution: differential privacy

In statistical analysis, the aim is to learn the information generated from a sample in order to learn about the structure of a population.

The study of the statistical database, however, must ensure the non-extraction of information about the individual in the sample, guaranteeing his privacy.

These two aspects do not always go together, and when information is used by an "adversary," an invidual's data could be breached.

The aim is to find a compromise between extracting possible information with the purpose of analysis and protecting the sensitive data of individuals.

The simplest solution that comes to mind is the complete anonymization of data, eliminating sensitive data, certainly guaranteeing the integrity of the individual but having no usefulness from a statistical point of view. The other is to have all information

available, with the advantage of conducting more precise statistical analysis but with the high danger of compromising the data released by a user.
To meet both of the above needs, two distinct categories were developed:

1: *Randomization*: involves the addition of an error term ("noise") while maintaining the qualities of the data from a statistical point of view;
2: *Generalization*: involves the formation of clusters based on similar features or attributes among the data.
The main topic of this analysis and solution of many problems, *differential privacy*, falls into the first group mentioned above.

1.3.1 The difference between privacy and differential privacy

In 1977, the mathematician Tore Dalenius, made a desideratum regarding statistical databases where he riches that: even with the presence of a collection of data, no additional information about the participant should be found than if it were not available. Confirmation that this request cannot be satisfied comes in 2008 from Cynthia Dwork, a computer scientist at Harvard University.
The researcher assures us that if a change in the database occurs, the result changes for both those who participate in the study and those who do not, compromising privacy. The crucial elements that can be harmful to the privacy of individuals are the data at the disposition of an "adversary."
She proves in her article [1] this assertion to us through a trivial example. What is taken into consideration is the height of a person, specifically a hypothetical subject named Terry Ross.
The database available initially contains the average heights of women of various nationalities and the "adversary" has as auxiliary information, that Terry Ross is two inches taller than the average Lithuanian woman.
If you have this data at your disposal, it is easy to come to know the height of your chosen subject, demonstrating how the same reasoning can be applied in many high

areas.

Dwork, highlights how it was not specified whether the individual is in the database or not, confirming the risk for even a subject not participating in the study. Differential privacy for this helps us by ensuring that participation in a study does not increase the likelihood of compromise.

This type of analysis aims not to learn anything more about the user than they know, but to get useful information about the entire population through the use of the data. Here we find the fundamental difference with privacy, which instead aims to ensure the non-use and protection of data released by the user, having shown in the previous examples that this is not possible in the presence of additional information.

Differential privacy can also guarantee high levels of user privacy, but always having the possibility to extract information from the database for further study purposes.

Thus, it is assured that with differential privacy, participation or non-participation in the study does not change the possibility of data release, as no action or non-action of the subject could avoid it.

1.3.2 The theory behind differential privacy

To formalise what is discussed above, the mathematical interpretation and the accompanying definition of *differential privacy* is expressed below taken from [2].

Before proceeding, a few essentials elements are presented in order to learn their meaning.

To ensure that "adversaries" do not violate the sensitive data of a study's "participants," or at least the probability is very low, we make explicit in three key points what has been learned so far.

Assume $X_1^n \in \mathcal{X}^n$ the input data sample and Z the result that is obtained from the chosen privacy mechanism.

i. Given otput Z, an "adversary" should not be able to recognize whether an individual is participating in a study, even if it knows everything (except one person);

ii. An individual's participation in multiple studies should not put his or her data at risk, but a good degree of privacy protection must be ensured; further analysis of a set of information is processed through a Markov chain $X_1^n \longrightarrow Z \longrightarrow Y$ reassuring the user on the security of his/her personal data;

iii. The privacy mechanism used $X_1^n \longrightarrow Z$, must be resistant to auxiliary information that may be known by an "adversary." If the information of a "participant" is present in X_1^n and the "adversary" is aware of other data, what it learns from the dataset should not give it any additional information.

Especially, someone can express the third point in Bayesian terms, assuming that the information that the "adversary" has available can be expressed as an a priori π distribution.
The knowledge of two different samples for only one individual, $\{x_1, ..., x_n\}$ and $\{x_1^{'}..., x_n^{'}\}$, by means of a private mechanism, does not lead to additional relevant knowledge a posteriori ($X_1^n \longrightarrow Z$).
The *differential privacy* is not an algorithm; it is a definition that assumes the use of a fundamental tool for protecting user information.
The model adds error, it also call it "noise" to the data to ensure the security of personal information, but while preserving the accuracy of the data set that is to be used.
The role of ϵ is decisive since it represents a metric of loss of privacy, i.e. the maximum distance that is measured between the output of sample x_1^n and that of y_1^n, also called the 'privacy parameter'.
If the value of ϵ is small, the two sample outputs will be very similar to each other, ensuring a high level of protection for users.

1.3.3 Application of differential privacy

The mechanism behind the *differential privacy* and the innovation introduced to fortify the protection of the population, was immediately successful and applied in several cases.

An example to mention is the company Apple, which in 2016, after the introduction of the new iOS 10 operating system, the company declared an application to all following operating systems, of differential privacy.

Following Apple are Google and Facebook as written in [14]. Thanks to this method, they They don't just guarantee a high level of protection for their users, but also the amount of data that can be used to identify specific trends or changes in user behaviour to improve their market.

In 2019, through the collaboration of Microsoft and the Institute for Quantitative Social Science at Harvard, a public data platform was developed ensuring differential privacy.

It was seen earlier how a user's data and information is collected in databases, which if not protected by a privacy mechanism can be easily hacked.

Microsoft explains how it is possible through the differential privacy mechanism to provide security for users.

In a privacy mechanism users send a *query*, i.e., a request, for data and the tools at the base respond by entering an error or "noise" .

This is done in order not to make the available data identifiable, returning an approximation of them as can be seen in Figure 1.3.

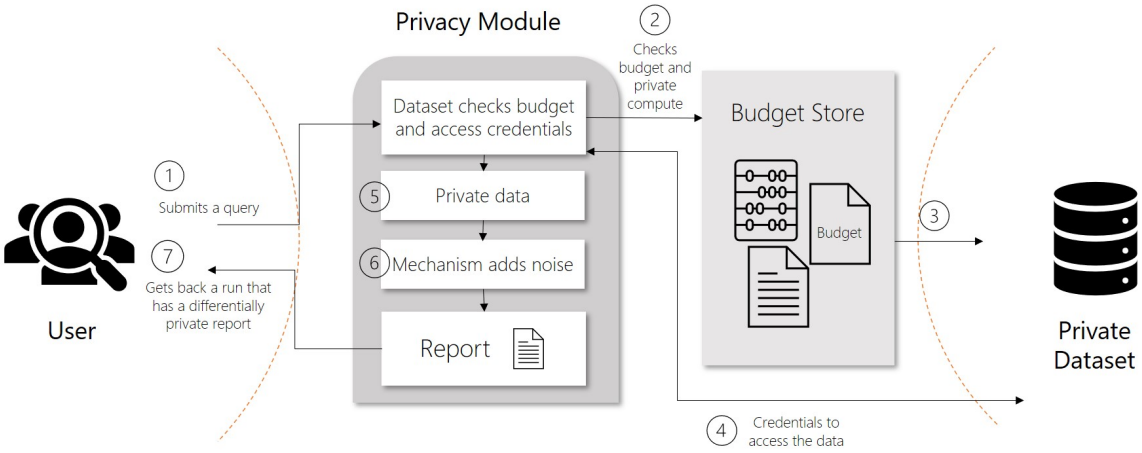


Figure 1.3: *Operation of Differential Privacy explained by Microsoft taken from [13]*

Chapter 2

Differential privacy tools

The problem expressed by Cynthia Dwork concerning additional information can be formalized with the mathematical definition of *differential privacy.*

Definition 2.0.1 (Neighbors dataset)**.** *Two different datasets $\mathcal{X}$ and $\mathcal{Y}$ are neighbors if they differ on at most one row, where each row corresponds to an individual.*
In this case the order of the rows is important.

Example 2.0.1. *We have 3 different datasets $\mathcal{X}, \mathcal{Y}, \mathcal{J}$ and we want to know if they are neighbours.*

Table 2.1: Example of three different datasets

X	Y	J
{1,1}	{0,0}	{1,1}
{0,0}	{0,0}	{1,1}
{1,1}	{1,1}	{0,0}

Looking the table, $\mathcal{X}$ and $\mathcal{Y}$ are neighbors, because they differ for only one row and $\mathcal{X}$ and $\mathcal{J}$ are not neighbors, because $\mathcal{J}$ contains the rows in a different order.

Theorem 2.0.2 (Markov kernel)**.** *Consider a probability space ($\Omega, \mathcal{B}(\Omega), \omega$), with Ω a sample space, $\mathcal{B}(\Omega)$ a σ -algebra of Borel and w a probability measure.*
A Markovian kernel is defined as, $\mathcal{Q} : \Omega \times \mathcal{B}(\Omega) \longrightarrow [0, 1]$, such that it is a measurable function in the first argument:

$$\mathcal{Q}(, \mathcal{A}) : \omega \longrightarrow \Omega, \forall \mathcal{A} \in \mathcal{B}(\Omega),$$

and a probability measure in the second,

$$\mathcal{Q}(q,) : \mathcal{B}(\Omega) \longrightarrow [0, 1], \forall q \in \Omega.$$

By construction the kernel defines a $Q : \Omega \longrightarrow \mathcal{B}(\Omega)$ map, where $\mathcal{B}(\Omega)$ is the space of probability measurements on Ω.

Definition 2.0.3 (Differencial privacy)**.** *[2] Let $\mathcal{Q}$ be a Markov kernel from $\mathcal{X}^n$ to an output $\mathcal{Z}$. $\mathcal{Q}$ is ϵ-differentially private if for all sets $\mathcal{S} \subset \mathcal{Z}$ and all sample $x_1^n \in \mathcal{X}^n$ and $x_1'^n \in \mathcal{X}^n$ differing in at most a single entry:*

$$\frac{Q(Z \in S | x_1^n)}{Q(Z \in S | x_1'^n)} \leq e^{\epsilon} \tag{2.1}$$

The above definition shows how data of a subject in a database, does not have a relevant effect on the result, so the absence of a user would not lead to large changes on the output.
Absence of a user would not lead to major changes in the output, with at most an error equal to e^{ϵ}.
The model of **Definition 2.0.4** is also called *centralized model*.
This name is given by the fact that the data samples are all controlled by a reference figure who manages the data, such as companies, institutions, hospitals or researchers.

Definition 2.0.4 (Local differencial privacy)**.** *[2] A Markov kernel $\mathcal{Q}$ from $\mathcal{X}$ to one output $\mathcal{Z}$ is ϵ-locally differencially private if for all measurable $\mathcal{S} \subset \mathcal{Z}$ and all x and $x' \in \mathcal{X}$,*

$$\frac{Q(Z \in S|x)}{Q(Z \in S|x')} \leq e^{\epsilon} \tag{2.2}$$

This model, called *local model*, differ from the Model (2.1) because this is stronger from a privacy point of view.

The two neighboring sample considered $\{x_1, x_2, ..., x_n\}$ and $\{x_1^{'}, x_2^{'}, ..., x_n^{'}\}$ differ by one observation, so the model in (2.2) consider the densities:

$$\frac{dQ(Z_1^n|x_i)}{dQ(Z_1^n|x_i^{'})} = \prod_{i=1}^{n} \frac{dQ(Z_i|x_i)}{dQ(Z_i|x_i^{'})} \leq e^{\epsilon}$$

where the difference are from the single entry where $x_i \neq x_i^{'}$.

The data providers, in this case, do not trust data collectors and make it protected before making it available to them.

The next Sections will present some of the mechanisms underlying differential privacy in the local or centralized models, where all these use the addition of an error to ensure privacy.

2.1 Randomized response

In the previous Chapter 1.3.2 we have introduced three assumptions from [2] in order to hypothesize the two models seen, which use an error to protect the data.

The oldest and also the most simple model based on this noise, is the *randomized response*, which was proposed by Warner in 1965. This model is a differential privacy technique that aims to reassure the user by introducing noise in relation to the original data.

[2]Assume that we want to know how many participants of a study have a have a certain characteristic, in this case if they have ever smoked. The portion of the population with this characteristic have a probability p, instead the others 1-p.

Obviously not all the population will answer the truth, so every single person have to flip a coin A. To achieve these results we refer to what John Duchi wrote in [6], the

probability $\mathrm{P}(x=1) = \frac{e^\epsilon}{1+e^\epsilon}$ correspond to the group 1, people that smoke or have tried to smoke, instead $\mathrm{P}(x=0) = \frac{1}{1+e^\epsilon}$ the group 0, people who have never tried to smoke. As support we will use a coin A where heads corresponds to group 1 and tails to group 0. With the aim of protecting the privacy of the participants, we ask to them to answer "Yes" if the coin corresponds to the group to which it belongs and "No" otherwise.
So we can write in this case that:

$$\frac{Q(Yes|x=0)}{Q(Yes|x=1)} = e^{-\epsilon}$$

and in the other case:

$$\frac{Q(No|x=0)}{Q(No|x=1)} = e^{\epsilon}$$

In general we can prove that this channel of randomize response guarantees ϵ- local privacy, in fact $\frac{Q(Z=z|x)}{Q(Z=z|x')} \in [e^{-\epsilon}, e^{\epsilon}]$ for all x, z.
So when there is differencial privacy we can't learn about a particular person in a set of data.

Example 2.1.1. *[18] Let us consider a single respondent in an analysis that collects answers about a sensitive topic. The respondent is instructed to flip a coin and if it is a head, he must answer the truth while if it is tails, then flip a second coin and respond "Yes" if heads and "No" if tails.*
The analysis show that Pr[Response ="Yes"/Truth = "Yes"] = $\frac{3}{4}$, because when the truth response is "Yes" there are two way to answer "Yes", the first one is when the first coin comes up tails with a probability $\frac{1}{2}$, the second one is when the first coin and the second come up heads, with a probability of $\frac{1}{2} * \frac{1}{2} = \frac{1}{4}$.
In the case of Pr[Response ="Yes"/Truth = "No"] = $\frac{1}{4}$, because there is only the case of the first coin comes up heads and second comes up tails. We can apply the similar reasoning to the case of a "No" answer and we obtain:

$$\frac{Pr[Response = "Yes"|Truth = "Yes"]}{Pr[Response = "Yes"|Truth = "No"]} = \frac{\frac{3}{4}}{\frac{1}{4}} = 3$$

$$\frac{Pr[Response = "No"|Truth = "No"]}{Pr[Response = "No"|Truth = "Yes"]} = \frac{\frac{3}{4}}{\frac{1}{4}} = 3$$

The randomized response described in this case is (ln 3, 0)-differentially private.

2.2 Generalised randomize response

We can extend the model based on random responses more generally.

Theorem 2.2.1 (Uniform Distribution)**.** *A continuous random variable Z that is an Uniform Distribution in a interval [a,b], represented by Z ∼ Uniform(a,b), have probability distribution function:*

$$f_Z(z) = \begin{cases} \frac{1}{b-a},\ a < z < b \\ 0,\ z < a\ or\ z > b \end{cases}$$

This mechanism is based on the output Z conditioned on x, a particular attribute chosen in a set of values $\{1, ..., k\}$.

$$Z|x = \begin{cases} x,\ \text{with}\ p = \frac{e^\epsilon}{k-1+e^\epsilon} \\ U(\frac{k}{x}),\ \text{with}\ p = \frac{k-1}{k-1+e^\epsilon} \end{cases}$$

where U is an Uniform distribution. The estimation of the true probability of the output Z, we can write $p_i = P(X = i)$ for a singular person, and the calculation is with the combination of the two probability reported above.
We obtain:

$$P(Z = i) = p_i \frac{e^\epsilon}{k-1+e^\epsilon} + (1-p_i)\frac{1}{k-1+e^\epsilon} = p_i \frac{e^\epsilon - 1}{e^\epsilon + k - 1} + \frac{1}{e^\epsilon + k - 1} \tag{2.3}$$

By definition $\sum_{j=1}^{k} P(Z = j) = 1$ because is a is a probability measure, and then to check the correctness of the result: $\sum_{j=1}^{k} P(Z = j) = \sum_{j=1}^{k} p_i \left(\frac{e^\epsilon - 1}{e^\epsilon + k - 1}\right) + \frac{k}{e^\epsilon + k - 1} = \left(\frac{e^\epsilon - 1}{e^\epsilon + k - 1}\right) \sum_{j=1}^{k} p_i + \frac{k}{e^\epsilon + k - 1} = 1$.
It's difficult to obtain the true probability then estimate. Suppose a sample of the

output Z of size n, $\hat{d}_n \in R_+^k$ and so the estimation of the true probability p is:

$$\hat{p}_n := \frac{e^\epsilon + k - 1}{e^\epsilon - 1}\left(\hat{d}_n - \frac{1}{e^\epsilon + k - 1}\mathbf{1}\right)$$

This result follows from the substitution of $P(Z = i) = \hat{d}_n$, $p_i = \hat{p}_n$ and reversal of the Equation (2.3), with $p^T\mathbf{1} = 1$.

Theorem 2.2.2 (Euclidean Norm)**.** *Assume a vector* $\mathbf{u} = (u_1, u_2, ..., u_n)$ *in a Euclidean n-space* $\mathcal{R}^n$*. The Euclidean Norm of* $\mathbf{u}$ *is defined by:*

$$||\boldsymbol{u}|| = \left(\sum_{j=1}^{n} u_j^2\right)^{\frac{1}{2}}$$

This probability satisfies the properties of non-distortion, that is $E[\hat{p}_n] = p$ and substituting the quantities in,

$$E[||\hat{p}_n - p||_2^2] = \left(\frac{e^\epsilon + k - 1}{e^\epsilon - 1}\right)^2 E[||\hat{d}_n - E[\hat{d}_n]||_2^2]$$

thanks to $E[\hat{d}_n] = \left(E[\hat{p}_n] + \frac{1}{e^\epsilon - 1}\mathbf{1}\right)\frac{e^\epsilon - 1}{e^\epsilon + k - 1} = \left(p + \frac{1}{e^\epsilon - 1}\mathbf{1}\right)\frac{e^\epsilon - 1}{e^\epsilon + k - 1}$ and applying the properties of the expected value.

Theorem 2.2.3 (Bernoulli distribution)**.** *A Bernoulli distribution is a discrete distribution, with only two outcomes,* $z = 1$ *with probability* p *for "success" and* $z = 0$ *with probability* $1 - p$ *for "failure". The probability density function is defined as:*

$$P_Z(z) = \begin{cases} p, \text{ for } z = 1 \\ 1 - p, \text{ for } z = 0 \end{cases}$$

Where $E[Z] = p$ *and* $V(Z) = p(1 - p)$.

Explaining the expected value formula and applying the variance formula of a Bernoulli distribution obtain the result:

$$E[||\hat{p}_n - p||_2^2] = \frac{1}{n}\left(\frac{e^\epsilon + k - 1}{e^\epsilon - 1}\right)^2 \sum_{j=1}^{k} P(Z = j)(1 - P(Z = j)) \tag{2.4}$$

We know that $\sum_{j=1}^{k} P(Z = j) = 1$, and so thanks to this result obtain that:

$$E[||\hat{p}_n - p||_2^2] \leq \frac{1}{n} \left(\frac{e^\epsilon + k - 1}{e^\epsilon - 1} \right)^2 .$$

To simplify the analysis, are considered the two cases where $\epsilon \leq 1$ and $\epsilon \geqslant \log k$. In the first c the mean l_2 square error has scale as $\frac{k^2}{n\epsilon^2}$ and the second one at worst $\frac{1}{n}$ which is the "non-private" mean square error. The generalized randomized response mechanism is ϵ - locally private, thanks to **Definition 2.0.4**.

2.3 The Laplace mechanism

The Laplace mechanism is used for the centralized model based on the addition of a Laplace noise for the exponetial tails, instead for the local differencial privacy is used the randomized response. For this method are used two tools: the Hamming metric, for the addiction of the noise and the Lipschitz costant, that is called also sensitivity.

Theorem 2.3.1 (Lipschitz constant)**.** *[2] The Lipschitz costant, based on some distance function, taking values in $\mathcal{R}^+$ and an order p>0, for two neighbourig databases x and $x^{'}$:*

$$Lip_{p,dist}(f) := sup \left\{ \frac{||f(x) - f(x^{'})||_p}{dist(x, x^{'})} | dist(x, x^{'}) > 0 \right\}$$

which is also called ***sensitivity****.*

Theorem 2.3.2 (Hamming metric)**.** *Assume two vectors $\mathbf{u} = (u_1, u_2, ..., u_n)$ and $\mathbf{v} = (v_1, v_2, ..., v_n)$, the Hamming distance d_{ham}:*

$$d_{ham}(u, v) = \sum_{i=1}^{n} \mathbf{1} \{u_i \neq v_i\}$$

that represents the number of differences between the two vectors $\mathbf{v}$ and $\mathbf{u}$.

Theorem 2.3.3 (Laplace distribution)**.** *The double exponential or Laplace distribution is a continuous distribution of two identical exponential distributions.*

The probability density function (PDF) is:

$$P(x) = \frac{1}{2b} e^{-|x-\mu|/b}$$

with $\mu \in \mathcal{R}$ *is the location parameter and* $b \in \mathcal{R}^+$ *is the scale parameter.*

$$D(x) = \frac{1}{2}\left[1 + sign(x-\mu)(1 - e^{-|x-\mu|/b})\right]$$

that is the integral of PDF and they have the same parameters. If $W \sim Laplace(b)$*, and* $W \in \mathcal{R}$*, then* $E[\mathcal{W}] = 0$ *because of symmetry, while* $E[\mathcal{W}^2] = \frac{1}{b}\int_0^\infty w^2 e^{-\frac{w}{b}} = 2b^2$ *applying the definition of second moment. Applying the Hamming metric to the definition of Lipschitz costant for a function* $f : \mathcal{X}^n \longrightarrow \mathcal{R}^d$ *with* $p = 1$ *we obtain:*

$$Lip_{1,d_{ham}}(f) = sup\left\{||f(x_1^n) - f(y_1^n)||_1 | d_{ham}(x_1^n, y_1^n) \leq 1\right\} \leq L \tag{2.5}$$

where L *is the Lipschitz constant.*

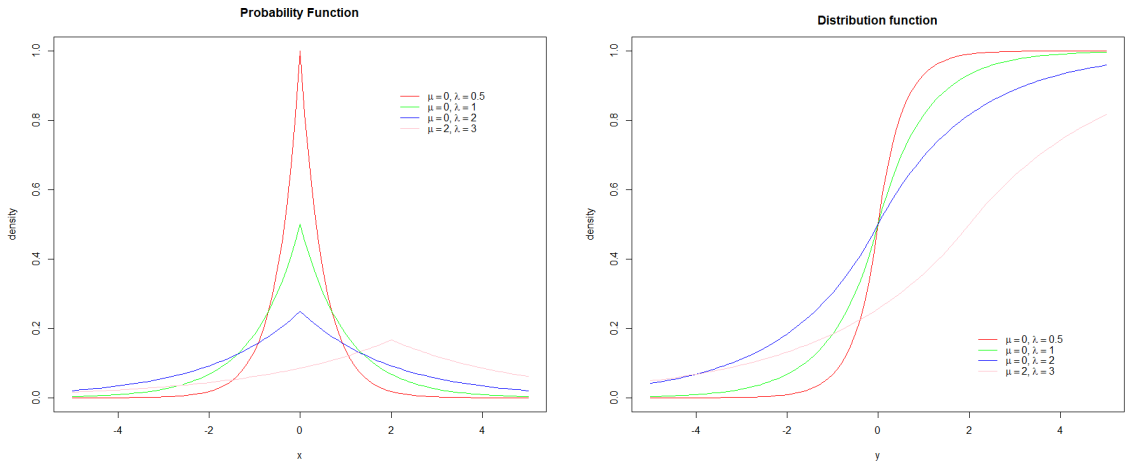


Figure 2.1: Density and distribution function of a Laplace distribution with different values of location parameter μ and scale parameter b.

Definition 2.3.4 (Laplace mechanism)**.** *[2] Let* $f\colon \mathcal{X}^n \longrightarrow \mathcal{R}^k$*. The Laplace mechanism is defined as:*

$$Z := f(X_1^n) + W$$

with a function $f(X_1^n)$ *and the addiction of* W *that is* $W_j \sim Laplace(\frac{L}{\epsilon})$ *i.i.d. .*

It must be demonstrated that also the Laplace mechanism Z is ϵ-differentially private taken from [7].

Theorem 2.3.5 (Triangle Inequality)**.** *Let* $\boldsymbol{x}$ *and* $\boldsymbol{y}$ *two different vectors. The triangle inequality is given by:*

$$|\boldsymbol{x}| - |\boldsymbol{y}| \leq |\boldsymbol{x} + \boldsymbol{y}| \leq |\boldsymbol{x}| + |\boldsymbol{y}|$$

Proof. Let consider two different dataset that differ for only one observation x and $x^{'} \in \mathcal{X}^n$ (neighbors database).
To prove differencial privacy it must be obtained that, the ratio of the density probability $p_{x'}(z)$ and $p_x(z)$ of the density functions of $Z(x)$ and $Z(x^{'})$ for a generic $z \in \mathcal{R}^k$, is $\frac{p_x(z)}{p_{x'}(z)} \leq e^{\epsilon}$.

$$\begin{aligned}
\frac{p_x(z)}{p_{x'}(z)} &= \frac{\prod_{i=1}^{k} exp(-\frac{\epsilon}{L}|f(x)_i - z_i|)}{\prod_{i=1}^{k} exp(-\frac{\epsilon}{L}|f(x^{'})_i - z_i|)} \\
&= \prod_{i=1}^{k} exp\Big(-\frac{\epsilon|f(x)_i - z_i|) - \epsilon||f(x^{'})_i - z_i|)}{L}\Big) \\
&\leq \prod_{i=1}^{k} exp\Big(-\frac{\epsilon(|f(x^{'})_i - f(x)_i|)}{L}\Big) \\
&= exp\Big(\frac{\epsilon \sum_{i=1}^{k}(|f(x)_i - f(x^{'})_i|)}{L}\Big) \\
&= exp\Big(\frac{\epsilon(||f(x) - f(x^{'})||_1)}{L}\Big) \\
&\leq exp(\epsilon)
\end{aligned}$$

thanks to the support of the triangle inequality for the third step. □

2.4 Laplace mechanism's application

[2] [7]The Laplace mechanism is used in different scenarios, in particular we will see below the use in estimation of means, histograms and counting queries.

2.4.1 Univariate mean estimation

The first application of Laplace mechanism is on univariate mean estimation, in particularis considered the **Definition 2.3.4.** for k = 1.
It is estimated $E[X]$, given a set of variables X, having values in a range [-c , c], with c a finite value ($c < \infty$).
It can be replaced the theoretical with the corresponding empirical estimator, $f(X_1^n) = \bar{X_n} = \frac{1}{n}\sum_{i=1}^{n} X_i$.
The Lipschitz costant in this case is $\frac{2c}{n}$, is tested in this in two different sample which only differ from one object, x and $x^{'} \in [-c, c]^n$:

$$|f(x) - f(x^{'})| = \frac{1}{n}|x_i - x_i^{'}| \leq \frac{2c}{n}$$

that respect the Hamming metric because $x_i \in [-c, c]$.
So the Laplace noise addiction mechanism is:

$$Z = \frac{1}{n}\sum_{i=1}^{n} X_i + W,\ W_i \sim Laplace\left(\frac{2c}{n\epsilon}\right)$$

The Laplace distribution is $W_i \sim Laplace(\frac{2c}{n\epsilon})$ with the variance:

$$V(W) = E[(Z - E[X])^2] = E[(\bar{X_n} - E[X])^2] + E[(Z - \bar{X_n})^2] \tag{2.6}$$

with $E[(\bar{X_n} - E[X])^2] = \frac{1}{n}Var(X)$ and $E[(Z - \bar{X_n})^2] = \frac{2(\frac{2c}{n})^2}{\epsilon^2} = \frac{8c^2}{n^2\epsilon^2}$ for the general result:

$$E[||Z - f(x_1^n)||_2^2] = \frac{2dL^2}{\epsilon^2}$$

The Equation (2.5) is written as:

$$E[(Z - E[X])^2] = \frac{1}{n}Var(X) + \frac{8c^2}{n^2\epsilon^2} \leq \frac{c^2}{n} + \frac{8c^2}{n^2\epsilon^2}$$

In this case the penalty is define as $\epsilon >> n^{-\frac{1}{2}}$ and with these values it can be defined that this mechanism is $\epsilon-$differentially private.

2.4.2 Counting Queries

Single Query

Another simple and similar application to the one just seen, is the counting queries. This problem is solved like the one just seen and responds to the request of how many subjects have a certain C characteristic.
In this case, every single person have a dichotomous variable $Y_i \in \{0, 1\}$, where 0 indicate that the person does not have the characteristic we are looking for and the opposite for 1.
Using the **Definition 2.3.4.** $f(X^n) = \sum_{i=1}^{n} Y_i$ and $W \sim Laplace\left(\frac{1}{\epsilon}\right)$, assuming Lipschitz costant equal to 1.
So the Laplace noise addiction mechanism is:

$$Z = \sum Y_i + W,\ W_i \sim Laplace\left(\frac{1}{\epsilon}\right)$$

Many Queries

In this second case one wants to answer to more queries.
Using **Definition 2.3.4.**, $f(X) = (f_1, f_2, ..., f_k)$, where each function answers a different question by counting the answers and assuming each individual counting query f_i has a sensitivity equal to 1.
Assuming two different dataset X and $X^{'}$ and we want to measure the L_1- sensitivity

with the Equation (2.5), that is:

$$L_1 = \sum_{i=1}^{k} |f_i(x) - f_j(y)| \leq \sum_{i}^{k} 1 = k$$

because some values cancel each other out.

So now, considering an overall sensitivity of k and so the Laplace random variables can be written like $W_i \sim Laplace(\frac{k}{\epsilon})$.

So the Laplace noise addiction mechanism is:

$$Z = \sum(f_1(X_i), ..., f_k(X_i)) + W,\ W_i \sim Laplace\left(\frac{k}{\epsilon}\right)$$

The difference between the single and many queries is the size of the error.

The first one has an order of error equal to $O(\frac{1}{\epsilon})$, so doesn't depend on the size of the database, instead in the second one each counting query has an error of $O(\frac{k}{\epsilon})$.

A criticality of the second approach, the counting of many queries, is that it uses a *non-adaptive* setting, must be introduced k counting queries at the beginning of the study and this is not optimal for analysis.

2.4.3 Histograms

The last estimation is another type of query, the *histogram query*.

In this case the queries are independent of each other, for example one wants to know how many people have an age X. This type of question is translated into function $f = (f_1, ..., f_k)$ where every function is the sum of subjects that have the age i.

In this case the sensitivity is equal to 2 because if they are considered two neighbouring datasets X and $X^{'}$, where x_j is replace by $x_j^{'}$, the counts change by 1 for both coordinates, j and $j^{'}$.

The Laplace mechanism has $f(X_i) = \sum_{j=1}^{n} I\{X_j = i\}$ for $i = 1, ..., k$ and $W_j \sim Laplace\left(\frac{2}{\epsilon}\right)$ i.i.d. variables and so:

$$Z = f(X) + W,\ W_j \sim Laplace\left(\frac{2}{\epsilon}\right)$$

Looking for the error, like before, have order $O\left(\frac{1}{\epsilon}\right)$, so it doesn't depend on "bins" k. An interesting thing is to measure the error for all counting variables, in particular for understand the accuracy of the Laplace mechanism one can start from the following result.

Observation 2.4.1. Assume an random variable Y $\sim$ Laplace (β) :

$$Pr[|Y| \geqslant t\beta] = \exp(-t)$$

Proof. If one integrate the Laplace PDF we get:

$$F_X : \begin{cases} \frac{1}{2}\exp\left(\frac{y-\mu}{b}\right) & \text{if } y < \mu \\ 1 - \frac{1}{2}\exp\left(-\frac{y-\mu}{b}\right) & \text{if } y \geq \mu \end{cases}$$

with $\mu = 0$ and $x = tb$ one have that:

$$\begin{aligned} Pr[|Y| \geqslant t\beta] &= 1 - Pr(-tb \leq Y \leq tb) = \\ &= 1 - (1 - \frac{1}{2}\exp(-t) - \frac{1}{2}\exp(-t)) = \\ &= \frac{2}{2}\exp(-t) = \exp(-t) \end{aligned}$$

then the **Observation 2.4.1** has been demonstrated. □

If one wants to knowthe accuracy of Laplace mechanism we consider a l_∞ error. From the **Observation 2.4.1.** and looking that $Y_i \sim Laplace(\frac{L}{\epsilon})$ have that:

$$\begin{aligned} Pr\left[||Y||_\infty \geqslant \ln\left(\left(\frac{t}{\beta}\right)\left(\frac{L}{\epsilon}\right)\right)\right] &= Pr\left[\max_{i\in[n]}|Y_i| \geqslant \ln\left(\left(\frac{t}{\beta}\right)\left(\frac{L}{\epsilon}\right)\right)\right] \\ &\leq n * Pr\left[|Y_i| \geqslant \ln\left(\left(\frac{t}{\beta}\right)\left(\frac{L}{\epsilon}\right)\right)\right] \\ &= n * \left(\frac{\beta}{n}\right) \\ &= \beta \end{aligned}$$

So the probability that any bin has error greater or equal to $2\frac{\ln(\frac{n}{\beta})}{\epsilon}$ is at most β.

2.5 Bayesian approach to information

2.5.1 Additional information and conditioning to output

The main point of the differential privacy, is to protect people from external attacks now and especially in the future.

If one consider the three point at the beginning of this paper (Section 1.3.2), in particular the second one $X^n \longrightarrow Z \longrightarrow Y$. That means that if there is an output Z, in a sample X^n, that is differentially private, then the underlying functions Y also satisfies this property. Considering two neighboring datasets, X and $X^{'}$, can be translated as the desire to satisfy privacy with a system of hypotheses which differ in a single entry in the alternative hypothesis: $H_0 : X_1^n = x_1^n$ and $H_1 : X_1^n = (x_1^{i-1}, x_i^{'}, x_{i+1}^n)$.

To test the hypothesis system we have to introduce a ϵ - *conditional hypothesis testing* for privacy of people:

$$Q(\Phi(Z) = 1|H_0, Z \in A) + Q(\Phi(Z) = 0|H_1, Z \in A) \geqslant 1 - \epsilon \tag{2.7}$$

where X^n is the input, Z is the output, $Q(A|H_0) > 0$ and $Q(A|H_1) > 0$.

This result says that the probability of failing a test is high, regardless of which hypothesis is true. To avoid this, we need to do one more hypothesis.

Observation 2.5.1. Assume Q a channel that is $\epsilon-$ differential private.

Then Q is also $\bar{\epsilon} = 1 - e^{-2\epsilon} \leq 2\epsilon$ - conditional hypothesis testing private.

Proof.

$$\begin{aligned}
Q(B|H_0, Z \in A) + Q(B^c|H_1, Z \in A) =& \frac{Q(A, B|H_0)}{Q(A|H_0)} + \frac{Q(A, B^c|H_1)}{Q(A|H_1)} \\
\geqslant& e^{-2\epsilon} \left(\frac{Q(A, B|H_1)}{Q(A|H_1)} + \frac{Q(A, B^c|H_1)}{Q(A|H_1)} \right) \\
\geqslant& e^{-2\epsilon} \left(\frac{Q(A, B|H_1) + Q(A, B^c|H_1)}{Q(A|H_1)} \right) \\
\geqslant& e^{-2\epsilon} \left(\frac{Q(A|H_1)}{Q(A|H_1)} \right) = e^{-2\epsilon}
\end{aligned}$$

where B= $\{z|\Phi(z) = 1\}$ is the region of acceptance of H_1, Φ is the test and $Q(A, B|H_1)+$ $Q(A, B^c|H_1) = Q(A|H_1)$. □

In this way, it was shown that even if you condition the channel on the output of the data set, you cannot recognise the initial data sample used.

2.5.2 Another method: Bayesian perspectives

[2]It can be considered and treated external information with another approach, using additional information more cautiously: the *Bayesian method*.
The fundamental tools for using this approach are a set of a priori distributions $\pi \in \mathcal{X}^n$ and the posterior distribution $\pi(.|Z)$ where Z is the output of a channel Q.

Theorem 2.5.1 (Conditional probability)**.** *Let A and B events with $P(B) > 0$, then the conditional probability of A given B is the number:*

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

Using the definition, it is easy to prove the following results, which will be useful for further analysis:

- *$P(A \cap B)$ = P(A/B)P(B);*
- *P(A) = P(A ∩B)+P(A ∩B^c) = P(A/B)P(B) + P(A/B^c)P(B^c);*
- *P(B/A) = $\frac{P(A|B)}{P(A)}P(B)$, (**Bayes' Rule**).*

Theorem 2.5.2 (Bayes Theorem)**.** *Let the events $B_1, ..., B_n$ and A from a partition of a sample space S.*
Given the event A we can write the total probability $P(A) = \sum_{i=1} P(A|B_i)P(B_i)$.
Using the Bayes' rule we have that:

$$P(B_i|A) = \frac{P(A|B_i)P(B_i)}{P(A)}$$

with $P(A) > 0$.

Theorem 2.5.3 (Bayesian posterior distribution)**.** *The Bayesian posterior distribution is defined via Bayes Theorem (**Theorem 2.5.2**), which states that given the likelihood function $l(x|\theta)$ of the sample x and a priori distribution on the parameter of interest, $\pi(\theta)$ then:*

$$\pi(\theta|x) = \frac{l(x|\theta)\pi(\theta)}{\int l(x|\theta)\pi(\theta)d\theta},$$

which can be written as

$$\pi(\theta|x)\alpha l(x|\theta)\pi(\theta)$$

because $\int l(x|\theta)\pi(\theta)d\theta$ is a constant that is not important to the result.

It is of interest to test a particular value x of a sample S.
With some assumptions you can show that what the adversary knows a posteriori is not so different from the initial result.
One of this assumptions is that, what an opponent knows about a subject in a dataset is independent of what he knows about the other components.
This can be mathematically translated as:

$$\pi(x_1^n) = \pi_{\left(x_1^{i-1}, x_{i+1}^n\right)}\pi_i(x_i), \tag{2.8}$$

With this result you can write the following statement:

Definition 2.5.4. *Let Q be an ϵ - differentially private channel and let π be a prior distribution with the characteristic presented in the Equation (2.8).*
The posterior density for X_i given z is:

$$e^{-\epsilon}\pi_i(x) \leq \pi(x|Z = z) \leq e^{\epsilon}\pi_i(x) \tag{2.9}$$

The **Definition 2.5.4.** says that even if an adversary has information and knowledge of a subject (prior distribution $\pi_i(x)$), the result conditional on the output does not change much (posterior distribution $\pi_i(x|Z = z)$).

Proof. To prove the Proposition we use the $\epsilon-$ differential privacy inequality and we

assume that w $\bigoplus_i$ x $= \left(x_1^{i-1}, x, x_{i+1}^n\right)$ so:

$$\begin{aligned}
\pi_i(x|Z=z) &= \frac{\int_{w\in\mathcal{X}^{n-1}} q(z|w\bigoplus_i x)\pi(w\bigoplus_i x)d\mu(w)}{\int_{w\in\mathcal{X}^{n-1}}\int_{x'\in\mathcal{X}} q(z|w\bigoplus_i x')\pi(w\bigoplus_i x')d\mu(w,x')} \\
&\leq e^{\epsilon}\frac{\int_{w\in\mathcal{X}^{n-1}} q(z|w\bigoplus_i x)\pi(w\bigoplus_i x)d\mu(w)}{\int_{w\in\mathcal{X}^{n-1}}\int_{x'\in\mathcal{X}} q(z|w\bigoplus_i x)\pi(w\bigoplus_i x')d\mu(w)d\mu(x')} \\
&= e^{\epsilon}\frac{\int_{w\in\mathcal{X}^{n-1}} q(z|w\bigoplus_i x)\pi_{\left(x_1^{i-1},x_{i+1}^n(w)\right)}d\mu(w)\pi_i(x)}{\int_{w\in\mathcal{X}^{n-1}}\pi_{\left(x_1^{i-1},x_{i+1}^n(w)\right)}d\mu(w)\int_{x'\in\mathcal{X}}\pi_i(x')d\mu(x')} \\
&= e^{\epsilon}\pi_i(x),
\end{aligned}$$

where μ is the base measure on $\mathcal{X}^{n-1}$ x $\mathcal{X}$ for the upper bound.

To prove the lower bound, proceed in the same way. □

Youcan prove that there are other types of prior and posterior for which ϵ - differential privacy holds.

If the adversary's prior follow a distribution that is invariant to permutation, the only thing to do is to change the Equation (2.9) to adapt it to the request made.

You can rewrite **Definition 2.5.4.** by taking up the result of **Theorem 2.5.1** (Bayes' Rule).

The result shows us the difference after observation of the output between the a priori and a posteriori distribution.

You have the following result:

Definition 2.5.5. *A channel $\mathcal{Q}$ from $\mathcal{X}^n \longrightarrow Z$ is $\epsilon-$ differentially private if and only if any prior distribution π on $\mathcal{X}^n$ and any observation $z \in \mathcal{Z}$, the posterior odds satisfy:*

$$\frac{\pi(x|z)}{\pi(x'|z)} \leq e^{\epsilon} \tag{2.10}$$

for all x and $x' \in \mathcal{X}^n$ with $d_{ham}(x,x') \leq 1$.

Proof. You know that $\pi(x|z) = \frac{q(z|x)\pi(x)}{q(z)}$, where q is the density of Z $\in \mathcal{Z}$.

Then:

$$\frac{\pi(x|z)}{\pi(x'|z)} = \frac{q(z|x)\pi(x)}{q(z|x')\pi(x')} \leq e^{\epsilon}\frac{\pi(x)}{\pi(x')}$$

for all $z, x, x^{'}$ if and only if Q is ϵ - differentially private.

□

It was seen earlier that, even in the Bayesian context, the a priori and a posteriori distribution do not change significantly for any Z output, for two datasets which differ by only one element, maintaining coherence with the definition of differential privacy.

Chapter 3

A weaker version of differential privacy

3.1 More real privacy definitions

In the previous Chapter, we saw how the definition of differential privacy is based on the definition of a noise with minimal dimensions.
Often the dimension of error does not reflect reality, which is why there is a need to make the definition of differential privacy more flexible.
What we work on is precisely the error, which is often larger than the one actually used, decreasing the amount of noise while trying to maintain a good degree of protection and accuracy.
In particular, we will see below two other types of privacy, reported in John Duchi's paper in [2]. These two privacy are based on the weakening of the error term leading to satisfactory data protection results which are: the (ϵ, δ)- differential privacy and (ϵ, α)-Rényi differential privacy.

3.1.1 (ϵ, δ)- differential privacy

The first type concerns very uncommon happenings, whose probability of occurrence is very low, i.e. in cases of catastrophic violations of privacy.

Definition 3.1.1 ((ϵ, δ)- differential privacy)**.** *[2] A channel from $\mathcal{X}^n$ to $\mathcal{Z}$, the output, is (ϵ, δ)-differentially private if for all sets $S \subset \mathcal{Z}$ and all neighbouring samples $x_1^n \in \mathcal{X}^n$*

and $y_1^n \in \mathcal{X}^n$*:*

$$Q(Z \in S|x_1^n) \leq e^{\epsilon} Q(Z \in S|y_1^n) + \delta \tag{3.1}$$

with $\epsilon, \delta \geqslant 0$.

In this privacy model, a new term is added δ, where the new interpretation of the Equation (3.1), is that the model is ϵ-privacy differentiated except for the additive term δ.

If $\delta = 0$ then the definition of ϵ-differential privacy and (ϵ, δ)- dirrential privacy coincide, but this event will not be considered in this paper.

[2]This model has the characteristic of decaying super-polynomially to zero and δ is that it satisfies the relation $\delta = \delta_n$ where $\delta_n << n^{-l}$ for any $l \in N$.

Even if this method has advantages, its use is not indicated as it is not appropriate for commonly used cases.

The term additive includes two cases where privacy is not given, considering the case of its neighbouring datasets, where in both cases ϵ-differential privacy is guaranteed with a probability of $1 - \delta$, and the difference is in the probability of the term δ.

The difference between the two failure modes can be significant.

In the first case, there is always some residual possibility of denial; in the second, the opponent occasionally learns the truth with certainty.

The probability of being wrong depends on the user's degree of acceptance of false positives, and depending on this, a good degree of protection is guaranteed or not.

A single privacy statement (ϵ, δ)-DP cannot distinguish between the two alternatives.

The (ϵ, δ)-differential privacy was initially thought of as a model to ensure privacy in the Gaussian mechanism.

Definition 3.1.2 (Gaussian mechanism)**.** *[6] The Gaussian mechanism* $\mathcal{G}$ *on the query function* $f : X \to \mathcal{R}^d$ *with sensitivity* L *applied over a database* $D \in X$ *outputs:*

$$\mathcal{G}(f(D)) = f(D) + \epsilon, \epsilon \sim N(0, \sigma^2 I_d) \tag{3.2}$$

where σ *denotes the standard deviation of the normal distribution* N *and is calibrated*

to the sensitivity L, I_d denotes the identity matrix with d diagonal elements.

Theorem 3.1.3 (Gaussian mechanism (ϵ, δ)-differential privacy)**.** *[22] A Gaussian error of a mechanism with $\sigma = L\sqrt{2\log(\frac{1.25}{\delta})}$ and $\epsilon, \delta \in (0, 1)$ is (ϵ, δ)-differentiated privacy.*

Gaussian mechanism cannot respect the (ϵ, δ)- differential privacy, because using this language there is always a possibility of privacy violation.
[2]A different definition of privacy closest to the analysis to do is based on Rényi divergences between distributions, monotonically transformed f divergences, where their structure will be simplified in these analyses.

3.1.2 A look at the f-divergences

To be able to define a natural relaxation of differential privacy, the Rény differential privacy, we must introduce the concept of divergence between two distributions.
To do so we will refer to the definitions given by John Duchi contents in [2].
There are different types of divergence measures and this type of privacy is based on *f-divergences* or Ali-Silvey divergence.
Let P and Q two probability distributions on a set $\mathcal{X}$, and we have a convex function with the distinction that $f(1) = 0$ such that $f : \mathcal{R}_+ \longrightarrow \mathcal{R}$.
We consider X a discrete set, then the f- *divergence* between P and Q is:

$$D_f(P||Q) := \sum_x k(x) f\left(\frac{p(x)}{q(x)}\right).$$

If we generalise the result and consider for each set X and a quantizer $\mathbf{q} : X \longrightarrow \{1, ..., m\}$ defined as $B_i = \mathbf{q}^{-1}(\{i\}) = \{x \in X |\ \mathbf{q}(x) = i\}$ be the portion the quantizer induces, we can define the quantized divergence:

$$D_f(P||Q|\mathbf{q}) = \sum_{i=1}^{m} Q(B_i) f\left(\frac{P(B_i)}{Q(B_i)}\right),$$

and in general the definition can be written as:

$$D_f(P||Q) := \sup\{D_f\left(P||Q|\mathbf{q}\right) \text{such that } \mathbf{q} \text{ quantizes X}\}.$$

If we instead consider two continuous distributions P and Q with respect to the base measure, as in our case of interest, we obtain:

$$D_f(P||Q) := \int_{\mathcal{X}} q(x) f\left(\frac{p(x)}{q(x)}\right) d\mu(x). \tag{3.3}$$

This divergence measure is fundamental to defining Rényi's differential privacy measure and subsequent analysis.

Definition 3.1.4 (Rényi-α-divergence)**.** *[2] Let P and Q be distributions on a space $\mathcal{X}$ with dentisies p and q compared to a measure μ and $\alpha \in [1, \infty]$.*
The Rényi-α-divergence between P and Q is:

$$D_\alpha(P||Q) := \frac{1}{\alpha - 1} \log \int \left(\frac{p(x)}{q(x)}\right)^\alpha q(x) d\mu(x). \tag{3.4}$$

The result (3.3) can be formulated as the expression (3.4) due to the following relationship:

$$D_f(P||Q) = \exp((\alpha - 1)) D_\alpha(P||Q)),$$

with $f(t) = t^\alpha - 1$.

3.2 Rényi-differential privacy

We describe below a generalisation of the definition of differential privacy based on Rény divergence defined in the Equation (3.4).

3.2.1 Tools for defining the measure of privacy

Before proceeding with the formal definition of Rényi-differential privacy, we must use additional supporting elements.
Ilya Mirnov, in his paper *"Rényi Differential Privacy"* (2017), demonstrates how the Rényi-α-divergence satisfies important properties to consider [9]:

1. Non negativity;
2. Monotonicity;
3. Probability preservation;
4. Weak triangle inequality.

Observation 3.2.1 (Non negativity)**.** [9] For two distributions P, Q and $\alpha \geqslant 1$:

$$D_\alpha(P||Q) \geqslant 0. \tag{3.5}$$

Observation 3.2.2 (Monotonicity)**.** [9] For two distributions P, Q and $1 \leq \alpha < \beta$:

$$D_\alpha(P||Q) \leq D_\beta(P||Q). \tag{3.6}$$

Observation 3.2.3 (Probability preservation)**.** [9] B is an event in $\mathcal{R}$, $\alpha > 1$ and two distributions P, Q in $\mathcal{R}$ defined on the same support:

$$P(B) \leq (\exp[D_\alpha(P||Q)]Q(B))^{(\alpha-1)/\alpha}. \tag{3.7}$$

Observation 3.2.4 (Weak triangle inequality)**.** [9] Let three distribution P, Q, R in $\mathcal{R}$ with $p, q > 1$ which satisfy $\frac{1}{p} + \frac{1}{q} = 1$ it is affirmed that

$$D_\alpha(P||Q) \leq \frac{\alpha - 1/p}{\alpha - 1} D_{p\alpha}(P||R) + D_{q(\alpha-1/p)}(R||Q). \tag{3.8}$$

with $\alpha > 1$.

Thanks to these tools, we can now proceed to the formal definition of differential privacy by Rényi.

Definition 3.2.1 (Rényi-differential privacy)**.** *[2] A channel $\mathcal{Q} : X^n \longrightarrow Z$ is (ϵ, α)-Rényi private if for all neighboring sample $x_1^n, y_1^n \in X^n$,*

$$D_\alpha(Q(|x_1^n)||Q(|y_1^n)) \leq \epsilon, \tag{3.9}$$

with $\epsilon \geqslant 0$ *and* $\alpha \in [1, \infty]$.

Remark 1. We can also define α for negative orders, but we will not consider this case.

In reference to (ϵ, δ)-differential privacy, Rényi differential privacy is a stronger type of privacy and each ϵ-differentiated privacy process is also (ϵ, α)-Rényi private.
In addition comparing pure differential and Rényi privacy many properties are in common.

3.2.2 The advantages of Rényi differential privacy

The privacy introduced satisfies several criteria necessary for it to be defined as a good data protection instrument.
In the following we will discuss about bad outcomes guarantee, robustness to auxiliary information, post processing and group privacy in the end, based on what Mirnov said in his writing [9].

Bad outcomes

In the first case, bad outcomes guarantee, there are some cases where the output of the mechanism is described as bad, this is why some people do not want to enter data into the database.
Differential privacy ensures that the probability of observing an erroneous result does not change whether a user's information is present in the dataset or not, the computer scientist says is valid also for the privacy of Rényi thanks to the Equation (3.7):

$$e^{-\epsilon} Q(Z \in S | y_1^n)^{\alpha/(\alpha-1)} \leq Q(Z \in S | x_1^n) \leq (e^{\epsilon} Q(Z \in S | y_1^n))^{(\alpha-1)/\alpha}$$

Post-processing

The researcher also says that an important benefit of this privacy is the fact that, if a function is Rényi differential privacy, the function's transformation also remains so. This ensures the security of the output even after the process is complete, being robust to manipulation.

Table 3.1: Comparison of differential and Rényi privacy properties taken from [9]

Differential Privacy	Rényi Differential Privacy	Property
$e^{-\epsilon} \leq \frac{Q(Z \in S \mid x_1^n)}{Q(Z \in S \mid y_1^n)} \leq e^{\epsilon}$	$e^{-\epsilon} Q(Z \in S \mid y_1^n)^{\frac{\alpha}{(\alpha-1)}} \leq Q(Z \in S \mid x_1^n) \leq (e^{\epsilon} Q(Z \in S \mid y_1^n))^{\frac{(\alpha-1)}{\alpha}}$	Change in probability of outcome Z
$\frac{R_{post}(x_1^n, y_1^n)}{R_{prior}(x_1^n, y_1^n)} \leq e^{\epsilon}$ always	$\mathrm{E}\left[\left\{\frac{R_{post}(x_1^n, y_1^n)}{R_{prior}(x_1^n, y_1^n)}\right\}^{\alpha-1}\right] \leq e^{\epsilon^{[(\alpha-1)\epsilon]}}$	Change in the Bayes' factor
$\lvert L \log R(x_1^n, y_1^n) \rvert \leq \epsilon$ always	E[Llog $R(x_1^n, y_1^n)] \leq \epsilon$	Change in log of Bayes' factor
f is ϵ-DP (or (α, ϵ)) $\longrightarrow g$ o f is ϵ-DP (or (α, ϵ)-RDP, resp.)		Post-processing
f, g is ϵ-DP (or (α, ϵ)) $\longrightarrow (f, g)$ is 3ϵ-DP (resp.,$(\alpha, 2\epsilon)$-RDP)		Adaptive sequential composition
f is ϵ-DP (or (α, ϵ)), g is 2^c-stable $\longrightarrow f$ o g is $2^c\epsilon$-DP (or $(\alpha/2^c, 3^c\epsilon)$-RDP)		Group privacy, pre-processing

Resistance to additional information

As pointed out in the previous Chapters, the introduction of differential privacy emerged in order to guarantee users security over their data even if an "adversary" has prior knowledge.

Mirnov assumes that he is in the conditions of **Definition 2.5.5.** of Chapter 2, $\frac{\pi(x|z)}{\pi(x^{'}|z)} = \frac{q(z|x)\pi(x)}{q(z|x^{'})\pi(x^{'})} \leq e^{\epsilon}\frac{\pi(x)}{\pi(x^{'})}$.

This result satisfies the fact that a mechanism under differential privacy does not change the input result except at the maximum of e^{ϵ}.

He states that if we talk about Rényi differential privacy, the protection provided by the mechanism is based on the variation of the posterior distribution.

Theorem 3.2.2 (Jensen's inequality)**.** *Let f a concave function and X a random variable,*

$$E[f(X)] \leq f(E(X)).$$

We assume that $R(x_1^n, y_1^n)$ is the Bayesian factor defined in **Theorem 2.5.2** and P= $f(x_1^n)$ and Q= $f(y_1^n)$, with an order α restricts the (α - 1)-th moment:

$$E_P\left[\left\{\frac{R_{post}(x_1^n, y_1^n)}{R_{prior}(x_1^n, y_1^n)}\right\}\right] = exp[(\alpha - 1)D_\alpha f(x_1^n)||f(y_1^n))].$$

Applying the logarithm on both sides and Jensen's disjointedness:

$$E_P\left[\log\left\{\frac{R_{post}(x_1^n, y_1^n)}{R_{prior}(x_1^n, y_1^n)}\right\}\right] \leq D_\alpha(f(x_1^n)||f(y_1^n)).$$

and consequently $logR_{post}(x_1^n, y_1^n) - logR_{prior}(x_1^n, y_1^n) \leq \epsilon$.

Preservation under conditions of adaptive sequencing

Another property that the scholar shows us is that, if two different function f and g are respectively ϵ_1- differentially private and ϵ_2- differentially private, then their sum is also $\epsilon_1 + \epsilon_2$- differentially private.

This result is also satisfied when choosing the function g based on the output of f, as

is shown in the following summary and will be important later to demonstrate other relevant results.

Theorem 3.2.3. *[9]Let $f : X \longrightarrow R_1$ be (α, ϵ_1)- Rényi differential privacy, than the mechanism defined as (K,Y), where $X \longrightarrow f(X)$ and $Y \longrightarrow g(K, X)$ satisfies $(\alpha, \epsilon_1 + \epsilon_2)$- Rényi differential privacy.*

Privacy group

So far, we have only considered pairs of neighbouring datasets, which are different for only one data item.
The IT specialist tells us how differential privacy can be defined in a weaker form but also considering datasets that are less close to each other.
In order to demonstrate this result, he takes up the concept of c-stable transformations.

Definition 3.2.4 (c-stable trasformation)**.** *[17]A trasformation $g(D)$ is c-stable if two neighbouring datasets D and $D^{'}$ satisfy:*

$$|g(D) \bigoplus g(D^{'})| \leq c \times |D \bigoplus D^{'}|.$$

Theorem 3.2.5. *[9]If $f : X \longrightarrow \mathcal{R}$ is (α, ϵ)- Rényi differential privacy, $g : X^{'} \longrightarrow X$ is 2^c-stable and $\alpha \geqslant 2^{c+1}$, the the composition of f o g is $(\frac{\alpha}{2}, 3^c\epsilon)$-Rényi differential privacy.*

3.2.3 Application of Rényi differential privacy

After looking at some characteristics of this type of privacy, examples are given concerning Gaussian mechanism, randomized response and Laplace distributions.
To do so we will refer to examples given by John Duchi [2] and Ilya Mironov in [9].

Gaussian mechanism

The first example is based on the divergence between two Gaussian distributions, in the second the result (2.5) is repeated considering l_2-norm.

Observation 3.2.5. [2] Let two Gaussian distribution defined with respectively mean μ_0, μ_1 and variance in common Σ:

$$D_\alpha(N(\mu_0, \Sigma)||N(\mu_1, \Sigma)) = \frac{\alpha}{2}(\mu_0 - \mu_1)^T \Sigma^{-1}(\mu_0 - \mu_1). \tag{3.10}$$

To find this result, it was used the forumla in Equation (3.4) and other supporting tools.

Proof. [2]Let p and q as the two densities of the normal distributions defined above and consider the integral $\int \left(\frac{p(x)}{q(x)}\right)^\alpha q(x)dx$ that is a part of the result (3.4).

$$\int \left(\frac{p(x)}{q(x)}\right)^\alpha q(x)dx = E_{\mu_1}\left[\frac{\exp(\frac{\alpha}{2}(X-\mu_1)^T\Sigma^{-1}(X-\mu_1))}{\exp(\frac{\alpha}{2}(X-\mu_0)^T\Sigma^{-1}(X-\mu_0))}\right]$$

and thanks to the property of powers can be rewritten as

$$\begin{aligned}
\int \left(\frac{p(x)}{q(x)}\right)^\alpha q(x)dx &= E_{\mu_1}\left[\exp\left(-\frac{\alpha}{2}(X-\mu_0)^T\Sigma^{-1}(X-\mu_0) + \frac{\alpha}{2}(X-\mu_1)^T\Sigma^{-1}(X-\mu_1)\right)\right] \\
&= E_{\mu_1}\left[\exp\left(-\frac{\alpha}{2}(\mu_0-\mu_1)^T\Sigma^{-1}(\mu_0-\mu_1) + \alpha(\mu_0-\mu_1)^T\Sigma^{-1}(X-\mu_1)\right)\right] \\
&= \exp\left(-\frac{\alpha}{2}(\mu_0-\mu_1)^T\Sigma^{-1}(\mu_0-\mu_1) + \frac{\alpha^2}{2}(\mu_0-\mu_1)^T\Sigma^{-1}(\mu_0-\mu_1)\right) \\
&= \exp\left(\left(-\frac{\alpha}{2} + \frac{\alpha^2}{2}\right)(\mu_0-\mu_1)^T\Sigma^{-1}(\mu_0-\mu_1)\right)
\end{aligned}$$

where second step was obtained by the relation $(X-\mu_0)^2 - (X-\mu_1)^2 = (\mu_0-\mu_1)^2 +$

$2(\mu_1 - \mu_0)(X - \mu_1)$, the third $(\mu_0 - \mu_1)^T\Sigma(X - \mu_1) \sim N(0, (\mu_1 - \mu_0)^T\Sigma^{-1}(\mu_1 - \mu_0))$ with $X(\mu_1, \Sigma)$ and the last step is to collect the common term $(\mu_0 - \mu_1)^T\Sigma^{-1}(\mu_0 - \mu_1)$. To find the result (3.10), apply the logarithm on both sides and divide by $\alpha - 1$. □

Example 3.2.1. *In this example we are in the same condition of the result (2.5) except that Hamming's metric d_{ham} uses l_2-norm:*

$$Z = f(X_1^n) + W, W \sim N(0, \sigma^2 I)$$

with Z the process defined with Gaussian errors, σ^2 positive and:

$$Lip_{2,d_{ham}}(f) = sup\{||f(x_1^n) - f(y_1^n)||_2 | d_{ham}(x_1^n, y_1^n) \leq 1\} \leq L$$

The measure of divergence between two normal distributions, J$\sim N(f(x), \sigma^2)$ and K$\sim N(f(x'), \sigma^2)$ is calculated as:

$$D_\alpha(J||K) = \frac{\alpha}{2\sigma^2}||f(x) - f(x')||_2^2 \leq \frac{\alpha}{2\sigma^2}L^2$$

Thus the mechanism Z satisfies the (ϵ, α)-Rényi differential privacy if and only if $\sigma^2 = \frac{L^2\alpha}{2\epsilon}$.

Randomized response

Now, as in the case of pure differential privacy, the random response model is considered (Section 2.1):

$$Z|x = \begin{cases} z, \text{ with } p \\ 1 - z, \text{ with } 1 - p \end{cases}$$

Using definitions **Definition 3.1.4.** and **Definition 3.2.1**, Mirnov arrives at the following result, (α, ϵ)-Rényi differential privacy, for $\alpha > 1$, with:

$$\epsilon = \frac{1}{\alpha - 1}\log(p^\alpha(1 - p)^{1-\alpha} + p^{1-\alpha}(1 - p)^\alpha).$$

Laplace mechanism

The last example concerns the Laplace mechanism as it was done in the Section 2.3. The mechanism is resumed as $f : \mathcal{X}^n \longrightarrow \mathcal{R}^k$. The Laplace mechanism is defined as:

$$Z := f(X_1^n) + W$$

with a function $f(X_1^n)$ and the addiction of W that is $W_j \sim Laplace(0, b)$ i.i.d., where b is the scale operator and 0 the position operator. To prove that the mechanism is Rényi privacy differential, the definition of divergence between two Laplace distributions is used (3.4) $D_\alpha(P||Q)$. $P \sim Laplace(0, b)$ and $Q \sim Laplace(1, b)$ are two Laplace distributions with densities equal to the result of the **Theorem 2.3.3**.
The integral of the Rényi divergence can be written as: $\frac{1}{\alpha-1} \log \int_R \left(\frac{p(x)}{q(x)}\right)^\alpha q(x) d\mu(x)$ where:

$$\left(\frac{p(x)}{q(x)}\right)^\alpha q(x) = \left(\frac{1}{2b} e^{-\frac{|x|}{b}}\right)^\alpha \left(\frac{1}{2b} e^{-\frac{|x-1|}{b}}\right)^{1-\alpha} \tag{3.11}$$

$$= \left(\frac{1}{2b}\right)^{\alpha+1-\alpha} \left(e^{\frac{-|x|\alpha}{b} + \frac{-|x-1|(\alpha-1)}{b}}\right) \tag{3.12}$$

The integral is now divided into three parts according to the value of x:

$$\begin{aligned}\int_{\mathcal{R}} \left(\frac{p(x)}{q(x)}\right)^\alpha q(x) d\mu(x) =& \int_{-\infty}^{0} \left(\frac{1}{2b}\right) \left(e^{\frac{x\alpha}{b} + \frac{(x-1)(\alpha-1)}{b}}\right) d\mu(x) \\ &+ \int_0^1 \left(\frac{1}{2b}\right) \left(e^{\frac{-x\alpha}{b} + \frac{(x-1)(\alpha-1)}{b}}\right) d\mu(x) + \int_{-\infty}^{0} \left(\frac{1}{2b}\right) \left(e^{\frac{-x\alpha}{b} - \frac{(x-1)(\alpha-1)}{b}}\right) d\mu(x) \\ =& \frac{1}{2} e^{\frac{(\alpha-1)}{b}} + \frac{1}{2(2\alpha-1)} (e^{\frac{(\alpha-1)}{b}} - e^{\frac{-\alpha}{b}}) + \frac{1}{2} e^{\frac{-\alpha}{b}}\end{aligned}$$

With this result it can now be written that:

$$D_\alpha(P||Q) = \frac{1}{\alpha-1} \log \left(\frac{\alpha}{2\alpha-1} e^{\frac{\alpha-1}{b}} + \frac{\alpha-1}{2\alpha-1} e^{\frac{-\alpha}{b}}\right) \tag{3.13}$$

and so the Laplace mechanism satisfies the $(\frac{1}{\alpha-1} \log(\frac{\alpha}{2\alpha-1} e^{\frac{\alpha-1}{b}} + \frac{\alpha-1}{2\alpha-1} e^{\frac{-\alpha}{b}}), \alpha)$- Rènyi differential privacy.

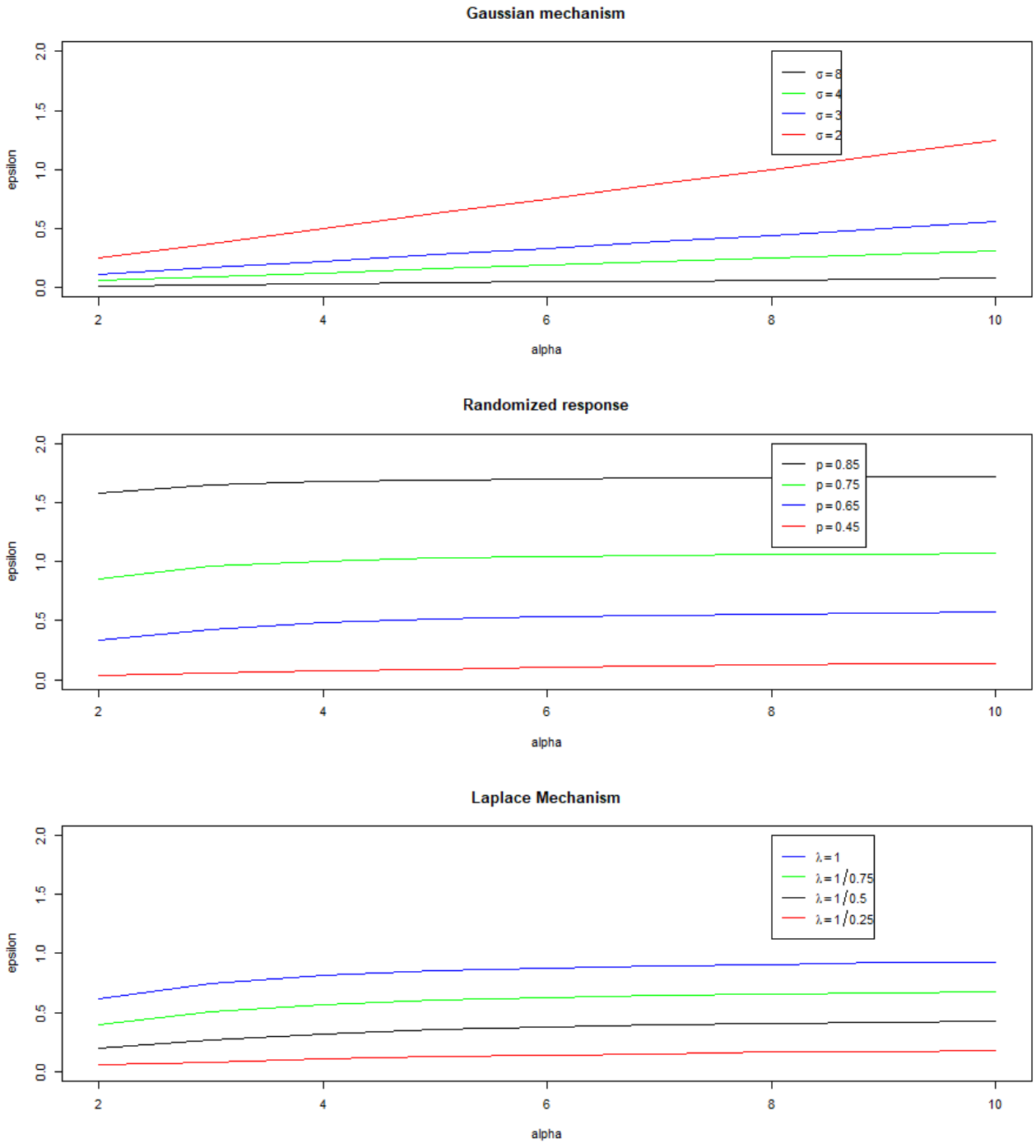


Figure 3.1: Comparison of mechanism with different parameters of (ϵ, α)- Rényi differential privacy using different parameter values such as [9]

.

From the above graph, it can be seen that the value of the error committed by the differentiated privacy mechanisms is always very small.
In particular, the Gaussian and Laplace mechanisms appear to have an error ϵ that is always less than 1, when the parameters vary.
The Table 3.2. compares and summarises the parameter values in the pure differential privacy and Rényi model.

Table 3.2: Comparison of parameters in differential and Rényi privacy taken from [9].

Differential Privacy	**Rényi Differential Privacy for α**	**Mechanism**
$\lvert\log \frac{p}{1-p}\rvert$	$\alpha > 1$: $\frac{1}{\alpha-1}\log(p^\alpha(1-p)^{1-\alpha} + (1-p)^\alpha p^{1-\alpha})$	Randomized Response
$\frac{1}{\lambda}$	$\alpha > 1$: $\frac{1}{\alpha-1}\log\left\{\frac{\alpha}{2\alpha-1}\exp(\frac{\alpha-1}{\lambda}) + \frac{\alpha-1}{2\alpha-1}\exp(-\frac{\alpha}{\lambda})\right\}$	Laplace Mechanism
∞	$\frac{\alpha}{2\sigma^2}$	Gaussian Mechanism

3.3 Relationship between different privacy measures

We have discussed and analysed the structure of three types of privacy which are: pure ϵ-differential privacy, (ϵ, δ)-differential privacy and lastly (α, ϵ)-Rényi differential privacy.
What we want to do now is to understand how and whether these measures are interconnected.
To proceed with the analysis we refer to the paper of John Duchi in [2].

3.3.1 ϵ-differential privacy and (ϵ, α)-Rényi differential

Proposition 3.3.1. [2] For $\alpha \geqslant 1$, $\epsilon \geqslant 0$ and $e^{-\epsilon} \leq \frac{P(B)}{Q(B)} \leq e^{\epsilon}$, with P, Q two distributions and B a generics set:

$$D_\alpha(P||Q) \leq \min\left\{\frac{3\alpha\epsilon^2}{2}, \epsilon\right\}$$

Theorem 3.3.2. *A system Q is $(min\{\frac{3\alpha\epsilon^2}{2}, \epsilon\}, \alpha)$-Rényi privacy differentiated if Q is assumed to be ϵ-privacy differentiated, with $\alpha \geqslant 1$ and $\epsilon \geqslant 0$.*

This theorem (thanks to **Proposition 3.3.1.**), confirms the fact that pure differential privacy is stronger than Rényi privacy, but from one the other can be obtained.

3.3.2 (ϵ, α)-Rényi differential privacy and (ϵ, δ)-differential privacy

Proposition 3.3.3. [2] For some collection B, with P and Q are $D_\alpha \leq \epsilon$,so:

$$P(B) \leq \exp\left(\frac{\alpha-1}{\alpha}\epsilon\right) Q(B)^{\frac{\alpha-1}{\alpha}}$$

for $\delta > 0$:

$$P(B) \leq \min\left\{\exp\left(\epsilon + \frac{1}{\alpha-1}\log\frac{1}{\delta}\right) Q(B), \delta\right\} \leq \exp\left(\epsilon + \frac{1}{\alpha-1}\log\frac{1}{\delta}\right) Q(B) + \delta$$

Theorem 3.3.4. *A system Q is $\left(\epsilon + \frac{1}{\alpha-1}\log\left(\frac{1}{\delta}\right), \delta\right)$-privacy differentiated if Q is assumed to be (ϵ, α)-Rényi privacy differentiated, with $\delta > 0$ and $\epsilon \geqslant 0$.*

This theorem (thanks to **Proposition 3.3.3.**), confirms the fact that Rényi differential privacy is stronger than (ϵ, δ)-differential privacy, but from one the other can be obtained.
What transpired in conclusion is the fact that each measure of privacy can be expressed as the others by varying the parameter values.
Starting with the strongest, as expected, namely pure ϵ- differential privacy up to the (ϵ, δ)- differential privacy.

3.4 Privacy and involvement in several studies

This section discusses the problem of how to manage privacy when a user participates in several studies also in the long term, one of the main problems of differential privacy introduced at the beginning of this paper.

It will be seen how Rényi differential privacy and the others privacy, as already mentioned in Section 3.2.2., have a fundamental property: *composition*.

3.4.1 The significance of composition

The composition, as mentioned above, is fundamental to solving the problem of a user's participation in multiple studies, but not only.
C. Dwork, G.N. Rothblum, S. Vadhan, in [16], explain the main applications of this property.

1. Composition is used on the same dataset with equal privacy mechanisms, ensuring a satisfactory level of privacy for users.;

2. Composition can also be used when there are different privacy models.
   If a user releases multiple data in different studies in which he or she participates protected by different types of privacy, this method provides good protection for his or her participants.

3. Composition is lastly used to move from simple privacy mechanisms to more complex ones.

In order to be able to explicate composition in the privacy mechanisms seen, it is necessary to introduce a couple of theorems.

Theorem 3.4.1 (Composition of ϵ- differential privacy mechanisms)**.** *[16]Assume a channel Q is ϵ- privacy differentiated, then the sum of k processes is $k\epsilon$- privacy differentiated under adaptive composition.*

Theorem 3.4.2 (Composition of (ϵ, δ)- differential privacy mechanisms)**.** *[16]Assume a channel Q is (ϵ, δ)- privacy differentiated, then the sum of k processes is $(k\epsilon, k\delta)$- privacy differentiated under adaptive composition.*

Thanks to these theorems, one can investigate behaviour in different types of privacy.

3.4.2 Composition for Rényi privacy

[2]The Rényi differential privacy ensures strong data protection, thus leading to a composition of different processes that are protected by privacy of different intensities. This can be proved by **Theorem 3.2.3.** concerning the property "preservation under conditions of adaptive sequencing".
[15]You have a sequence of output $\{Z_1(I_1, X), ..., Z_k(I_K, X)\}$ where Z_k is (ϵ, α) - Rényi differential privacy ,X is a a set of data, I_k regarding the k-th mechanism with a different input, depending on the previous outputs and inputs $\{(Z(I_j, X), I_j) : j < k\}$.

Theorem 3.4.3. *[15]Let (ϵ, α) - Rényi differential privacy, $\epsilon_i < \infty$ and $\alpha \geqslant 1$ with $i = 1, .., k$ then $(M_1, ..., Mk)$ is $(\sum_{k=1}^{K} \epsilon_k, \alpha)$- Rényi differential privacy.*

The theorems defined above can be formalised by using a privacy algorithm that guarantees that private channel composition remains as it is.

Algorithm 1 Privacy algorithm and composition taken from [2]

Require: Family of channels $\mathcal{Q}$ $b \in \{0, 1\}$.
1: **for** $k = 1, 2, ...$ **do**
2: Adversary choose space $\mathcal{X}, n \in \mathcal{N}$ and two neighbouring datasets $X, X' \in X^n$.
3: Adversary chooses private channel $\mathcal{Q}_k \in \mathcal{Q}$.
4: Adversary observes one sample $Z_k \sim Q_k(|x^{(b)})$.
5: **end for**

The algorithm reported by John Duchi, [2] basically works with an "adversary" that chooses an arbitrary space and two neighbouring datasets, meaning they differ by one entry.
Once the two datasets are chosen, they are privatised and the "adversary" can repeat the process iteratively.
At the end of the algorithm, one looks at the sequence of mechanisms chosen by the adversary and wonders if it is privatised.
This result is ensured thanks to the **Theorem 3.4.1, 3.4.2, 3.4.3**.

3.4.3 Composition for (ϵ, δ)- differential privacy

We have just seen how Rényi's measure of privacy enjoys the property of composition. As seen in Section 3.3. with the results expressed by John Duchi, namely the connection of the three privacy measures seen, different privacy mechanisms can be derived through additional composition between processes.
We now speak in terms of a more advanced composition, starting with a pure differentiated privacy mechanism, which thanks to the composition can be expressed in terms of (ϵ, δ)-differential privacy.

Theorem 3.4.4. *[2]Let Q a channel ϵ- differentiated privacy, and then a composition of k channels is $k\epsilon$- differentially private. The additional composition of the channel sequence is:*

$$\left(\frac{3k}{2}\epsilon^2 + \sqrt{6k\log\frac{1}{\delta}}, \delta\right)$$

differentially private, for $\delta > 0$ which is the additive term.

Proof. [2] Thanks to **Theorem 3.4.1.** it can be stated what is written in the first part of the theorem.
To continue the demonstration, we refer to the **Theorem 3.3.1.** and **3.3.2.** which respectively link pure differential privacy to Rényi privacy and Rényi privacy to (ϵ, δ)-differential privacy.
The first pass to the privacy of Rényi has an error of $(\frac{3\alpha}{2}\epsilon^2)$ while the second step $\epsilon + \frac{1}{\alpha - 1\log\frac{1}{\delta}}$.
In conclusion, the composition of k channels lead to a process $(\frac{3k}{2}\epsilon^2 + \frac{3k\gamma}{2}\epsilon^2 + \frac{1}{\gamma}\log\frac{1}{\delta}, \delta)$-differentially private where $\gamma = 1 + \alpha > 0$ and $\delta > 0$ and optimises relative to γ.

□

Chapter 4

Model application and validation

In this chapter we discuss some examples seen in the course of this paper and some new from a practical point of view.

For the implementation and viewing of the results, was used the Rstudio software for the entire analysis.

In recent years, differential privacy has been widely applied to protect data before it is disseminated, adding noise to information. Many researchers remain sceptical about the use of this method, as they are afraid that the error entered will compromise the quality of the data and consequently the analyses do not reflect reality.

This type of analysis, however, guarantees the secrecy of user data while maintaining a good degree of accuracy with a low margin of error.

A relevant example is found in the U.S. with the census, whose data are essential to be able to predict the behavior of the population, number of inhabitants, age groups, resources to allocate and other relevant information.

On the other hand, however, the censuses allow many sensitive data concerning the individual and his/her private life, thus having to protect them.

Differential privacy helps in this, where in fact in the census of U.S. in 2020, compared to 2010, it was chosen to introduce it in order to protect the safety of the participants but not compromising fundamental analyses any more.

4.1 Three mechanisms compared

Differential privacy is applied to three basic mechanisms considering different artificially created datasets for analysis.

The privatisation process follows what has been said before: given an unprotected process in a given dataset, one wants to guarantee security to a user.

This can be done by adding noise, with a specific distribution, to prevent the identification of a subject in the dataset.

The first model considered is the *Laplace mechanism* which is defined as in **Definition 2.3.4.** and assumes that the error follows a Laplace distribution.

The resulting process Z is privacy differentiated defined as $Z := f(D) + W$ with a function f and W that is $W_j \sim Laplace\left(\frac{L}{\epsilon}\right)$ with identical and independent distributions with scale equal to the ratio of sensitivity L and ϵ.

The second mechanism analysed is the *Gaussian mechanism* as defined in **Definition 3.1.2.** of Chapter 3, which works in the same way as the Laplace mechanism with the difference that the added noise follows a Gaussian distribution.

The mechanism is defined as $Z := f(D) + W$, $W \sim N(0, \sigma^2 I_d)$ with f a function, and σ denotes the standard deviation of the normal distribution with mean equal to zero.

The last mechanism is not been discussed until now and is called *Bernstein mechanism*, in summary this mechanism is based on approximation via the Bernstein polynomials and the addition of noise with a Laplace distribution.

This process is explained and developed in the following sections thanks to [21].

For the analysis, the mean of a sample was taken as the function for the process to be differentiated, but this can be replaced by many others.

4.1.1 Preliminary analysis

To perform the analysis, we use the package **"diffpriv"** introduced by Benjamin I. P. Rubinstein and Francesco Alda in [19] implemented in Rstudio.

This package makes statistical methods, through different functions, from non-private

to privacy-differentiated and introduces a tool for estimating the sensitivity of processes according to the type of function defined.

The first thing to choose is the dataset, which, as we reported earlier, is constructed manually and varies in numerosity and composition depending on the model to be used.

The Laplace and Gaussian mechanisms guarantee a differentiated privacy process for two datasets D and $D^{'}$, which are neighbouring, meaning that they differ by only one input.

In this analysis, pure ϵ-differential privacy is used for the Laplace and Bernstein mechanism, while for the Gaussian mechanism we use (ϵ, δ)-differential privacy.

The parameters of these two types of pravacy are contained in the **diffpriv** package in the respective classes **DPParamsEps** (ϵ) and **DPParamsDel**(ϵ, δ).

Consequently, the function of the library **releaseResponse** takes as input the function to be privatised f, the parameters defined above ϵ, δ and the data set D.

The responses are protected by differential privacy depending on the distribution chosen for the noise in **DPMech**, taken from **diffpriv**.

The individual functions are then discussed in more detail according to the chosen error distribution.

We now go into the specifics of the mechanisms and talk about how the analysis will take place in each of them and which functions and libraries will be used.

4.2 Laplace mechanism and ϵ-differential privacy

4.2.1 Tools

For the Laplace mechanism, three manually constructed datasets of different sizes are used and compared.

The first D, being a matrix of size 1000x1000, the second D_1 100x100 and the third D_2 10x10, randomly generated by a Uniform distribution with values between 0 and 1.

We have chosen as the function to privatise f the mean of a sample, such as in the

example in Section 2.4.1.

For this analysis, reference is made to the definition of pure ϵ-differential privacy in Chapter 2 (**Definition 2.0.3**).

4.2.2 Algorithm

The purpose of this and subsequent algorithms, is to privatise the responses of functions not covered by a privacy system of a given dataset.

The idea behind this algorithm follows the line introduced in the theory.

Taken two different datasets that differ by only one entry D and $D^{'} \in \mathcal{X}^n$, a mechanism $Z : \mathcal{X}^n \longrightarrow \mathcal{R}$ guarantees differential privacy between two datasets.

This result means that the addition or removal of an individual to a dataset is not privacy-relevant, while maintaining the same level of protection.

In mathematical terms, the mechanism must satisfy the relationship in **Definition 2.0.3**, leading a non-privatised $f : \mathcal{X}^n \longrightarrow \mathcal{B}$ function to be differential privacy by the addition of a noise W from a Laplace distribution .

This distribution has as location parameter zero and as scale the ratio between the sensitivity L and the value of ϵ.

The sensitivity of the process is calculated by L_1 norm considering the two neighbouring datasets D and $D^{'}$ and the non-privacy differentiated function f, $L = \sum_{i=1}^{k} |f_i(x) - f_j(y)|$.

To obtain the Z process $Z := f(D) + W$ covered by the ϵ-privacy differential, the function **releaseResponse** was used. This function takes as input the privacy parameters, in this case ϵ contained in the class **DPParamsEps**, dataset D and a target function f. The formal translation of what we have said can be written as $D = [0,1]^n$ and each neighbouring dataset $D^{'} = [0,1]^n$, $f(D) = \bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$, $L = sup\{||f(D) - f(D^{'})||_1\}$ $= sup\{\frac{1}{n}|\sum_{i=1}^{n} D_i - \sum_{i=1}^{n} D_i^{'}|\} = \frac{1}{n}$ from the relation: $\frac{1}{n}|D - D^{'}| \leq \frac{1}{n}$.

The mechanism $Z := f(D) + W$ with W that is $W_j \sim Laplace\left(\frac{L}{\epsilon}\right)$ i.i.d. is ϵ-differentiated privacy.

Algorithm 2 Laplace mechanism

Require: Dataset $D \in [0,1]^n$, privacy parameters ϵ, function f
Ensure: A numeric private response of the mechanism Z
1: Calculate l_1 norm for sensitivity $L = |f(D) - f(D')|$.
2: Generating errors W from a Laplace distribution, $W \sim Laplace\left(\frac{L}{\epsilon}\right)$.
3: Private response, $Z = f(D) + W$.

4.2.3 Results of Laplace mechanism

In this analysis we consider different values of ϵ for the three different datasets considered, to understand how the behaviour of the privacy mechanism varies depending on the value of the chosen error.

In particular, we study the process with different values of $\epsilon = 1$, $\epsilon = 2$ and $\epsilon = 5$ for different datasets.

We will also look at the difference between privatised and non-privatised answers with different chosen values and dataset sizes.

The choice of these three values was made in order to see how the behaviour of the process changes in terms of accuracy.

What can be observed from the below tables, particularly the last one Table 4.4, is that the difference between the non-privatised and privatised function is almost non-existent even with the introduction of an error term following a Laplace distribution.

The numerosity of the dataset is decisive for the order of magnitude of the error, as expected, in fact for a numerosity of $n = 1000000$ you get up to an order of magnitude of 10^{-8} for $\epsilon = 5$ while with n=100 you have 10^{-3} for $\epsilon = 1$.

As much data as the mechanism has, better it can estimate the process, getting as close as possible to the true value.

Therefore, the size of the dataset you have is crucial to ensure the accuracy of the data as well as its protection.

As far as the value of ϵ is considered, it can be seen from the data in the tables that if the value of ϵ is high, the results are more accurate but the level of privacy is lower.

On the other hand, if ϵ is small, the level of privacy is high but the results are less accurate, and this can be seen especially for the largest dataset $n = 1000000$.

The decision to choose the value of ϵ is up to the individual, evaluating the trade-off between accuracy and data protection.

Table 4.1: Comparison of privatised and non-privatised mechanisms with different numerosity and $\epsilon = 1$.

$\epsilon = 1$	Private response	Non - private response $f(D)$
n = 1000000	4.995×10^{-1}	4.995×10^{-1}
n = 10000	5.013×10^{-1}	5.014×10^{-1}
n = 100	5.221×10^{-1}	5.234×10^{-1}

Table 4.2: Comparison of privatised and non-privatised mechanisms with different numerosity and $\epsilon = 2$.

$\epsilon = 2$	Private response	Non - private response $f(D)$
n = 1000000	4.995×10^{-1}	4.995×10^{-1}
n = 10000	5.106×10^{-1}	5.015×10^{-1}
n = 100	5.254×10^{-1}	5.237×10^{-1}

Table 4.3: Comparison of privatised and non-privatised mechanisms with different numerosity and $\epsilon = 5$.

$\epsilon = 5$	Private response	Non - private response $f(D)$
n = 1000000	4.995×10^{-1}	4.995×10^{-1}
n = 10000	5.014×10^{-1}	5.015×10^{-1}
n = 100	5.243×10^{-1}	5.234×10^{-1}

Table 4.4: Difference in absolute terms between privatised and non-privatised mechanisms when varying numerosity and ϵ.

|private - $f(D)$|	$\epsilon = 1$	$\epsilon = 2$	$\epsilon = 5$
n = 1000000	1.224×10^{-6}	2.491×10^{-7}	8.226×10^{-8}
n = 10000	8.775×10^{-5}	7.532×10^{-6}	2.236×10^{-5}
n = 100	1.305×10^{-3}	1.992×10^{-3}	9.017×10^{-4}

4.3 Gaussian mechanism and (ϵ, δ)-differential privacy

4.3.1 Tools

The example we are now going to present refers to a differentiated privacy mechanism characterised by the introduction of an error with a Gaussian distribution.
Three manually constructed datasets of different sizes are used and compared.
The first D being a matrix of size 100x50, the second D_1 10x50 and the third D_2 10x10, randomly generated by a Uniform distribution with values between -0.5 and 0.5.
In terms of privacy we consider a weaker type of privacy, already introduced in Chapter 3, namely the (ϵ, δ)-differential privacy, where the general definition can be found in **Theorem 3.1.1**. This type of privacy provides ϵ-privacy protection for the mechanism except for a probability of δ, in this case equal to 10^{-4}.

4.3.2 Algorithm

As in the algorithm used to privatise a function via the Laplace mechanism, here too an algorithm is constructed based on what has been introduced theoretically.
Taken two different datasets that differ by only one entry D and $D^{'} \in \mathcal{X}^n$, a mechanism $Z : \mathcal{X}^n \longrightarrow \mathcal{R}$ guarantees (ϵ, δ)-differential privacy between two datasets. This result means that the addition or removal of an individual to a dataset is not privacy-relevant, while maintaining the same level of protection.
In mathematical terms, the mechanism must satisfy the relationship in **Definition 3.1.1.**, leading an non-privatised $f : \mathcal{X}^n \longrightarrow \mathcal{B}$ function to be (ϵ, δ)-differential privacy.
The differentiated privacy mechanism is obtained by adding the target function to a noise W from a Gaussian distribution.
This distribution has as a mean equal to zero and a standard deviation equal to $\sigma = \frac{L\sqrt{2\log\left(\frac{1.25}{\delta}\right)}}{\epsilon}$, which due to this ensures (ϵ, δ)- differential privacy to our initial function.
The sensitivity of the process, in this case, is calculated by L_2 sensitivity considering the two neighbouring datasets D and $D^{'}$ and the non-privacy differentiated function

f, $L = (\sum_{i=1}^{k}(f_i(x) - f_j(y))^2)^{1/2}$.

To obtain the Z process $Z := f(D) + W$ covered by the (ϵ, δ)-privacy differential, the function **releaseResponse** was used. This function takes as input the privacy parameters, in this case ϵ and δ contained in the class **DPParamsDel**, the target function f and dataset D.

The formal translation of what we have said can be written as $D = [-0.5, 0.5]^n$ and each neighbouring dataset $D^{'} = [-0.5, 0.5]^n$, $f(D) = \bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$, $L_2 = ||f(D) - f(D^{'})||_2 = \frac{\sqrt{d}}{n}$.

The mechanism $Z := f(D) + W$ with W that is $W_j \sim Normal\,(0, \sigma^2)$ i.i.d. is (ϵ, δ)-differentiated privacy, with d the number of differences between the two samples and a fixed $\delta = 10^{-4}$.

Algorithm 3 Gaussian mechanism

Require: Dataset $D \in [-0.5, 0.5]^n$, privacy parameters (δ, ϵ), target function f
Ensure: A numeric private response of the mechanism Z

1: Calculate l_2 sensitivity $L = ||f(D) - f(D^{'})||_2$.
2: Generating errors W from a Gaussian distribution $\sigma = \frac{L\sqrt{2\log\left(\frac{1.25}{\delta}\right)}}{\epsilon}$, $W \sim N\,(0, \sigma^2)$.
3: Private response $Z = f(D) + W$.

4.3.3 Results of Gaussian mechanism

Table 4.5: Comparison of privatised and non-privatised mechanisms with different numerosity and $\epsilon = 1$, global sensivity $L = \frac{1}{n}$ and $\delta = 10^{-4}$.

$\epsilon = 1$, $\delta = 10^{-4}$	Private response	Non - private response $f(D)$
n = 5000	1.738×10^{-3}	1.213×10^{-4}
n = 500	4.047×10^{-3}	2.059×10^{-2}
n = 100	-3.190×10^{-2}	-1.763×10^{-2}

Table 4.6: Comparison of privatised and non-privatised mechanisms with different numerosity and $\epsilon = 3$, global sensivity $L = \frac{1}{n}$ and $\delta = 10^{-4}$.

$\epsilon = 3$, $\delta = 10^{-4}$	Private response	Non - private response $f(D)$
n = 5000	-6.912×10^{-5}	1.213×10^{-4}
n = 500	2.451×10^{-2}	2.059×10^{-2}
n = 100	-1.992×10^{-2}	-1.763×10^{-2}

Table 4.7: Difference in absolute terms between privatised and non-privatised mechanisms when varying numerosity and $\epsilon = 1$ and 3, global sensivity $\frac{1}{n}$ and $\delta = 10^{-4}$.

|private - $f(D)$|	$\epsilon = 1$	$\epsilon = 3$
n = 5000	1.617×10^{-3}	3.073×10^{-4}
n = 500	1.654×10^{-2}	3.917×10^{-3}
n = 100	1.427×10^{-2}	2.292×10^{-3}

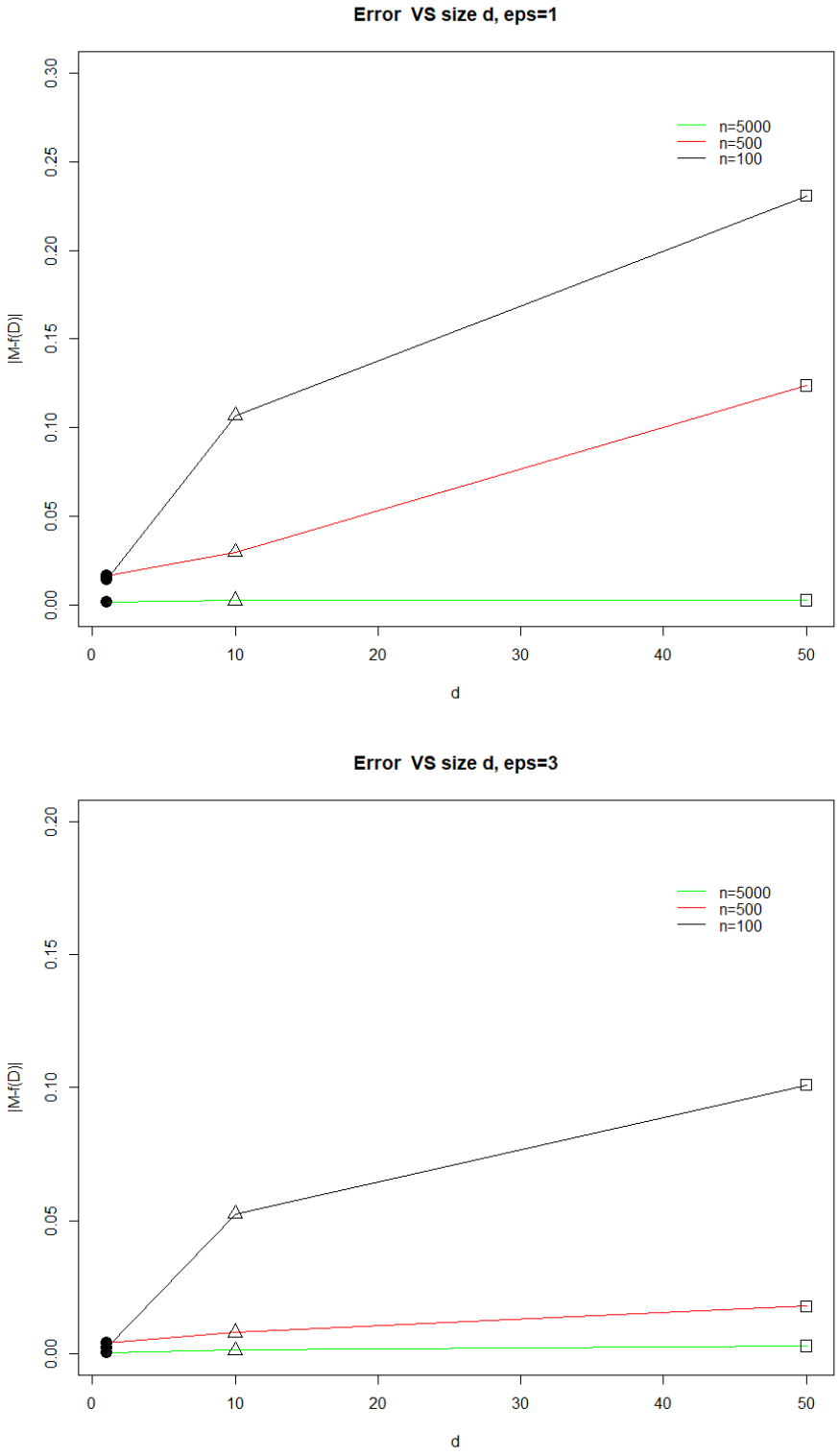


Figure 4.1: Comparison of the difference between the differentiated privacy mechanism and the non-privatised function, with a value of $\epsilon = 1$, $\epsilon = 3$ and $\delta = 10^{-4}$ with different dataset sizes when varying d.

In this analysis we consider different values of ϵ and fixed $\delta = 10^{-4}$ for three different datasets considered, to understand how the behaviour of the privacy mechanism varies depending on the value of the error.

In particular, we study the process with different values of $\epsilon = 1$ and $\epsilon = 3$ for all the analysis.

We will also look at the difference between privatised and non-privatised answers with different dimension of d for sensitivity and dataset sizes.

The choice of these values was made in order to see how the behaviour of the process changes in terms of accuracy. As in the case of Laplace's distribution, even with the introduction of an error following a Normal distribution, the difference between the true non-privatised and privatised result is minimal despite having relaxed the definition of privacy.

Also in this case the size of the dataset is decisive for the order of magnitude of the error, as expected, in fact for a numerosity of $n = 5000$ you get up to an order of magnitude equal to 10^{-4} for $\epsilon = 3$ while with n=100 you have 10^{-2} for $\epsilon = 1$.

In contrast to the previous example, with the introduction of δ, the difference between the results is greater, but confirming the fact that the size of the dataset is decisive in ensuring greater protection.

As far as the value of ϵ is considered, it can be seen from the data in the tables that if the value of ϵ is high, the results are more accurate but the level of privacy is lower.

On the other hand, if ϵ is small, the level of privacy is high but the results are less accurate, and this can be seen especially for the largest dataset $n = 5000$.

One can also observe in the graphs in Figure 4.1. how the difference between the privatised and non-privatised process varies as d changes, the value of d is the numerator of the overall sensitivity of the Gaussian mechanism $L_2 = \frac{\sqrt{d}}{n}$.

It is observed that as d increases, the error between the privatised and non-privatised process increases because the two datasets are more distant and obviously for larger ϵ the order of magnitude is smaller. The decision to choose the value of ϵ is up to the individual, evaluating the trade-off between accuracy and data protection, and also for the value of δ in this case equal to 10^{-4}.

4.4 The Bernstein mechanism and ϵ-differential privacy

Different from the previous mechanisms, this process has not yet been seen in this paper.

This mechanism, called *Bernstein Mechanism* was introduced by F.Aldà and B.I.P. Rubinstein in their paper [21] and is based on the approximation by Bernstein basis polynomials.

Definition 4.4.1 (Bernstein basis polynomials)**.** *[23]The Bernstein basis polynomials of degree n are defined for $i \in \{0, n\}$ by:*

$$B_{i,n}(t) = \binom{n}{i} t^i (1-t)^{n-i} \tag{4.1}$$

with $\binom{n}{i} = \frac{n!}{i!(n-i)!}$ and $B_{i,n}(t) = 0$, if $i < 0$ or $i > n$.

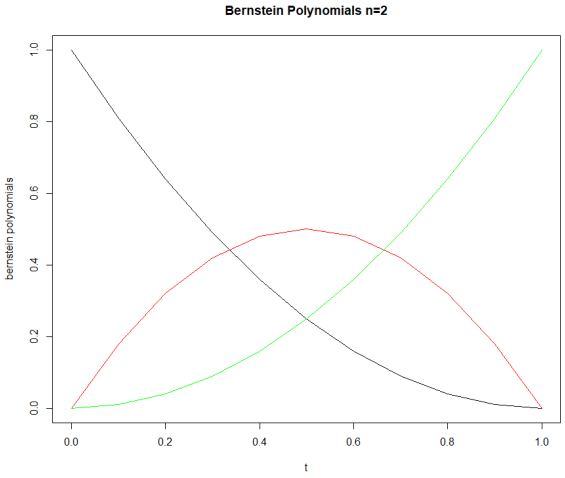


Figure 4.2: Example of Bernstein polynomials with degree $n = 2$

.

Proposition 4.4.2. [21] For $t \in [0, 1]$, for any integer n and $0 \leq n$ we have that $B_{i,n} \geqslant 0$ and $\sum_{i=0}^{n} B_{i,n}(t) = 1$.

Definition 4.4.3 (Bernstein polynomials)**.** *[21]Let $f : [0, 1] \longrightarrow \mathcal{R}$ and an integer n,*

Bernstein polynomials of degree n are defined for $i \in \{0, n\}$ *by:*

$$P_n(f;t) = \sum_{i=0}^{n} f\left(\frac{i}{n}\right) B_{i,n}(t) \tag{4.2}$$

It is now illustrated how Aldà and Rubistein ensured differential privacy for the Bernstein mechanism, but to proceeded to the main theorem, two further supporting tools must be introduced.

Definition 4.4.4 (Smooth function)**.** *[21]Let* $h, l > 0$ *integer,* $T > 0$ *and a function* $f : [0,1]^l \longrightarrow \mathcal{R}$*. The function* f *is* (h, T)*-smooth if* $C^h([0,1]^l)$ *and its partial derivatives up to the order* h *are all bounded by* T*.*

Definition 4.4.5 (Hölder continuous function)**.** *[21] Let* $0 \leq \gamma \leq 1$ *and* $L > 0$*. We have for every* $x, x^{'} \in [0,1]^l$*,* $|f(x) - f(x^{'})| \leq L||x - y||_{\infty}^{\gamma}$ *and so* f *is* (γ, L)*-Hölder continuous.*

With these tools, one can now define two differential privacy theorems for the Bernstein mechanism in [21].

Theorem 4.4.6 (ϵ-differential privacy in Bernstein mechanism)**.** *Let* L *and* $T > 0$ *costants,* $0 \leq \gamma \leq 1$*,* $l, h \in \{0, 1, ..., n\}$ *and* $F : \mathcal{X}^n \times \mathcal{Y} \longrightarrow \mathcal{R}$ *the target function. For* $\epsilon > 0$ *the Bernstein mechanism* Z *is* ϵ*-differentiated privacy.*

Theorem 4.4.7 ((ϵ, δ)-differential privacy in Bernstein mechanism)**.** *Let* L *and* $T > 0$ *costants,* $0 \leq \gamma \leq 1$*,* $l, h \in \{0, 1, ..., n\}$ *and* $F : \mathcal{X}^n \times \mathcal{Y} \longrightarrow \mathcal{R}$ *the target function. For* $\epsilon > 0$ *the Bernstein mechanism* Z*, with the perturbation scale equal to* $\lambda_\delta = \frac{2L_F\sqrt{2(n+1)^l l \log(\frac{1}{\delta})}}{\epsilon}$ *is* (ϵ, δ)*-differentiated privacy.*

4.4.1 Tools

The following is one of the examples given in the document [21] by applying it in practice and following the track in the document. Before proceeding, some supporting definitions are outlined.

Theorem 4.4.8 (Priestly-Chao regression)**.** *[25] Let $y_i = m(x_i) + e_i$ a non parametric regression model with y_i = response variable, $m(x_i)$ = unknown smooth function and e_i = error $\sim N(0, \sigma^2)$ for $i = 1, .., N$ then:*

$$\hat{m}(x_i) = \frac{\delta}{h} \sum_{i=1}^{N} K\left(\frac{x - x_i}{h}\right) Y_i, \tag{4.3}$$

with $\delta = \frac{(b-a)}{n}$, $x \in (a, b)$, h the kernel smoothness and K the kernel.

Three models are used for this analysis, namely: a Priestly-Chao regression with a Gaussian kernel (Equation 4.3.), its approximation by the non-privatised Bernstein polynomials (Equation 4.2) and a privatised regression using Bernstein's method.
Three manually constructed datasets of different sizes are used and compared, $D \in \mathcal{X}^n$. The first D being a matrix of size 100x2, the second D_1 500x2 and the third D_2 5000x2, randomly generated by a Uniform distribution for covariates x and and dependent variables through the equation $y = f(x) + e$ with $f(x) = sin(x * 10) * x$ and $e \sim N(0, 0.04)$.
In terms of privacy we consider the pure ϵ-differential privacy as in the case of Laplace mechanism.

4.4.2 Algorithm

Berstein's mechanism, like the two previous ones, guarantees differential privacy coverage of a generic function $F : \mathcal{X}^n \times \mathcal{Y} \longrightarrow \mathcal{R}$.
The mechanism uses the Bernstein polynomial, defined in Equation 4.2., iterated for the target function F, with a coverage K.
The differentiated privacy mechanism Z is obtained with addition of an error following Laplace distribution, with λ perturbation scale.
This is possible thanks to the function **DPMechBernstein** for the K iterated target function F and **releaseResponse** restoring the privatised model Z.
In contrast to the previous examples, in this case the value of the sensitivity L_F is not known and we need to use the function **sensitivitySampler** implemented and explained

in the document [24] and whose and this procedure is reported in the **Algorithm 4**.

Algorithm 4 Sensitivity Sampler

Require: f target function, distribution P, size of the dataset N, level of confidence γ, sample size m.
Ensure: Global sensitivity $L = G_k$.
1: **for** $i = 1$ **to** m **do**
2: Sample $D \sim P^{n+1}$
3: Set $G_i = ||f(D_{1...n}) - f(D_{1...n-1,n+1})||_B$
4: **end for**
5: Sort $G_1, ..., G_m$ as $G_{(1)} \leq ... \leq G_{(m)}$

In this specific example $m = 500$ that is, 500 random pairs are taken from the dataset from sample P and confidence $\gamma = 0.2$.

The idea of this algorithm is that independent and identically distributed observations $G_1, ...G_n$ are generated from the data sample P and thanks to these observations, it is possible to calculate the sensitivity of the differentiated privacy process.

If the cumulative distribution function (CDF) of the observations is not known, then the Uniform approximation is resorted to using the empirical CDF on the available sample.

The operation of Bernstein's general mechanism is presented in detail in **Algorithm 5**. The idea under this process is that the privatised process Z is obtain, through the sum of the function F adding the noise which follows a Laplace distribution with a λ parameter, over a coverage F identifying the possible datasets.

The parameter $\lambda = \frac{L_F(n+1)^l}{\epsilon}$ called perturbation scale, depends on the sensitivity L_F, the sample size n and the value of ϵ.

Algorithm 5 The Bernstein mechanism

Require: Dataset $D \in \mathcal{X}^n$, sensitivity L, target function F, parameters n, h, ϵ.
Ensure: Differentiated privacy function $\{Z(p)|p \in P\}$.
1: Lattice cover of $\mathcal{Y}$ $P \leftarrow (\{0, \frac{1}{n}, ..., 1\})^l$
2: Perturbation scale $\lambda \leftarrow \frac{L(n+1)^l}{\epsilon}$
3: **for** $p = (p_1, ..., p_l) \in P$ **do**
4: $Z = F(D) + W$, where $W \sim Laplace(\lambda)$ i.i.d.
5: **end for**

4.4.3 Results of the Bernstein mechanism

In this analysis we consider different values of ϵ for three different size of datasets and three different model.

In particular, we study the process with different values of $\epsilon = 1$ and $\epsilon = 3$ for all the analysis. The choice of these values was made in order to see how the behaviour of the process changes in terms of accuracy.

In contrast to the models differentiated by Laplace and Gaussian distribution, the Bernstein mechanism has values of the differentiated process somewhat more distant from those of the target function, as can be seen in the Tables 4.8 and 4.9.

The mean square error MSE $= \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$ was calculated with y_i equal to target function and Bernstein polynomial approssimation and with $\hat{y}_i$ equal to Bernstein function privatized.

The difference, of course, is more pronounced when the value of ϵ is smaller and when the size of the dataset is lower.

This is evident in the case of the Figure 4.3 where the scale of the variable Y shrinks considerably from N = 100 to N = 5000 and the estimated functions are closer and closer to each other as the numerosity of the dataset increases. From the results of the tables, a small difference can also be seen when varying the value of ϵ showing that the difference between the two results is greater when ϵ is smaller.

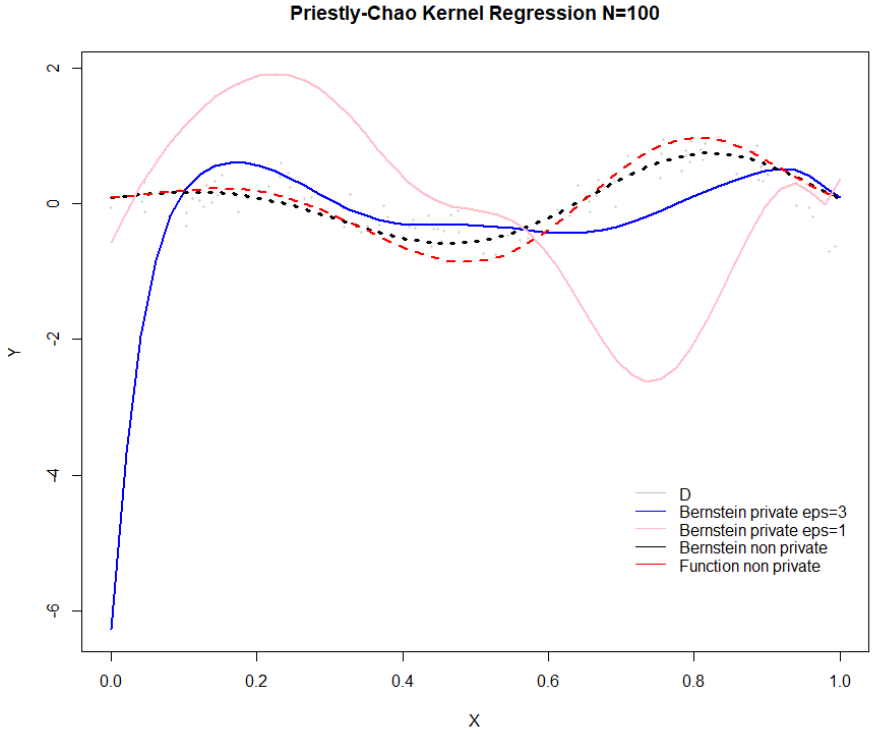
Another aspect that is apparent from the figure is that when the size of the dataset is large, the privatised Bernstein function is practically coincident with the non-privatised Bernstein approximation.

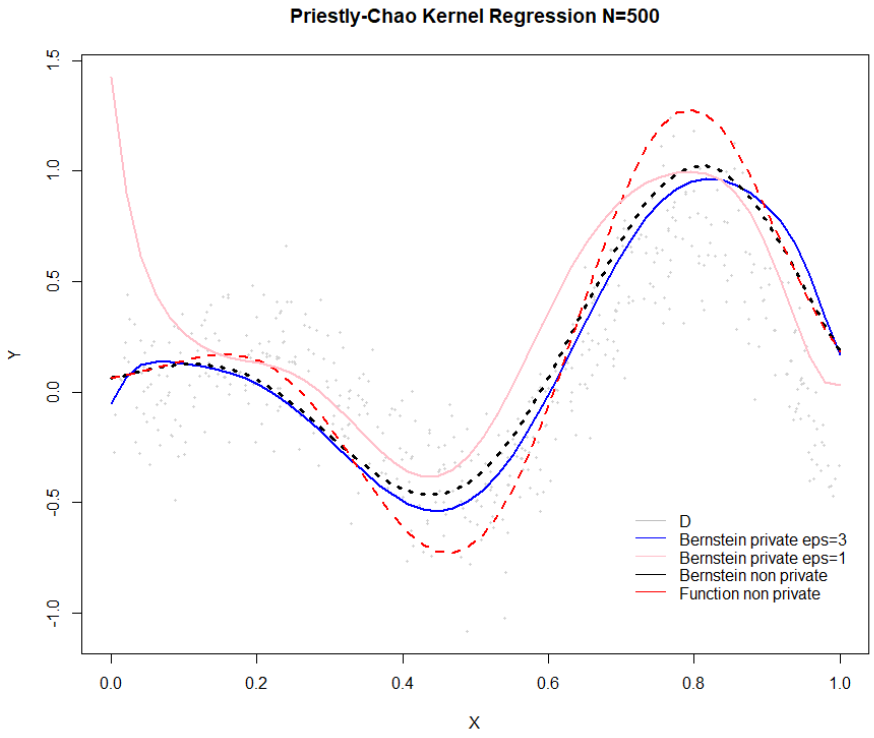
Table 4.8: MSE, with y the target function and the Bernstein approximation and $\hat{y}$ the function privatised through the Bernstein mechanism with $\epsilon = 1$.

MSE $\epsilon = 1$	Target function	Bernstein approximation
n = 5000	2.359×10^{-2}	6.806×10^{-4}
n = 500	1.165×10^{-1}	8.258×10^{-2}
n = 100	2.605	2.392

Table 4.9: MSE, with y the target function and the Bernstein approximation and $\hat{y}$ the function privatised through the Bernstein mechanism with $\epsilon = 3$.

MSE $\epsilon = 3$	Target function	Bernstein approximation
n = 5000	2.360×10^{-4}	4.075×10^{-5}
n = 500	2.445×10^{-2}	3.601×10^{-3}
n = 100	1.409	1.337


Priestly-Chao Kernel Regression N=100
Y
X
2
0
-2
-4
-6
0.0
0.2
0.4
0.6
0.8
1.0
D
Bernstein private eps=3
Bernstein private eps=1
Bernstein non private
Function non private


Priestly-Chao Kernel Regression N=500
Y
X
1.5
1.0
0.5
0.0
-0.5
-1.0
0.0
0.2
0.4
0.6
0.8
1.0
D
Bernstein private eps=3
Bernstein private eps=1
Bernstein non private
Function non private

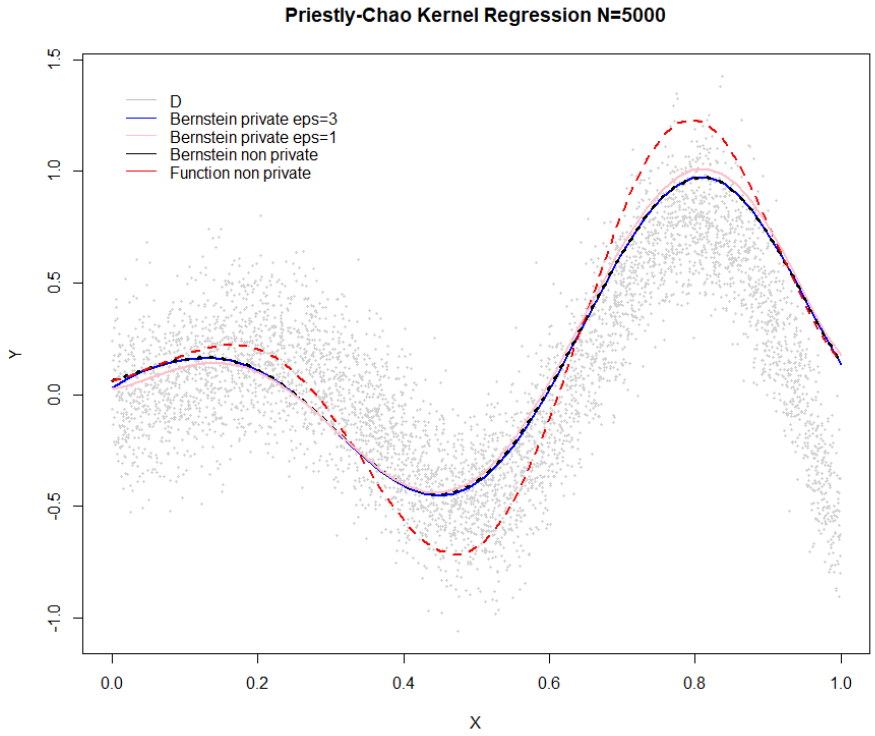


Figure 4.3: Comparison between three model of the non-privatised target function, the non-privatised Bernstein approximation and the privatised Bernstein mechanism with $\epsilon = 1$ and $\epsilon = 3$, with three different size of datasets $N = 100, 500, 5000$.

4.5 Conclusion of the analysis and observations

To conclude and summarise the analysis so far, it is evident that the choice of the parameter ϵ and the parameter δ in the case of the differential privacy (ϵ, δ) is decisive for the desired result, as their values can lead to very different scenarios.

In addition, the other important aspect is the amount of data available in the database, which, due to its size, can lead to greater adherence to reality and precision in the analysis.

Comparing the three models, it can be seen that there is not a great difference in the results of the three processes, all having satisfactory results, with a greater divergence in Bernstein's mechanism.

All three models are good processes for ensuring a good level of data protection while maintaining accuracy.

The first two mechanisms, Laplace's and Gaussian, are simpler to deal with and fewer parameters are involved. Bernstein's mechanism is more complex in terms of the algorithm and the number of parameters involved, but is less restrictive in terms of the functions that can be used.

Further analyses could be done, such as determining the optimal value of the parameter ϵ and δ on the basis of the available data, evaluating different target functions by the privatisation mechanism or lastly introducing more complex models.

In addition, a mechanism could be implemented to provide a model with (ϵ, α)-Rényi differentiated privacy protection, or through the relationships seen in the Section 3.3., one could derive the different types of privacy connected to each other.

Conclusions

Aim of this thesis is to demonstrate, apply and understand whether differential privacy could be a good solution for the protection of a community of people who release personal information while still being able to consider the data available as accurate.
In order to arrive at an answer, we first wanted to conduct a theoretical investigation of the topic and then applied the mechanisms and types of privacy we considered most relevant within the thesis.
Pure differential privacy and local privacy provide a very good level of data protection but lose some of their accuracy, thus leading to their use being less useful as they are unreliable.
The (ϵ, δ)-differential privacy or better (ϵ, α)-Rényi differential privacy has demonstrated a good adherence to reality and versatility thanks to the introduction of a new way of defining privacy based on Rènyi divergence and to the properties it enjoys.
The latter type of privacy in fact demonstrated a good level of protection while maintaining data accuracy.
It has been proven how differential privacy, which protects large volumes of data from big companies, is constructed simply by using elementary basic statistical distributions such as the Laplace, the Normal, the Uniform or the Exponential.
What in fact emerges from the data is that, especially with pure differential privacy, the difference in the quantity of the dataset and the level of ϵ, indicating greater or lesser data protection, makes a large difference in the accuracy of the data, leading this type of mechanism to be less attractive to companies and researchers.
This is one of the reasons why organisations often do not have or do not want to implement this type of privacy, as it requires a substantial financial commitment in terms

of resources or personnel that is not always accessible to all.

Another limitation is the fact that often not all companies are ready to sacrifice data accuracy for the protection of data, reporting to use differential privacy with perhaps so much ϵ level as to have no relevance in terms of security.

One aspect to be worked on in the future is the possibility of keeping the level of privacy from decreasing as the number of compositions, the presence of a user in several databases.

This can be solved by combining other types of data protection such as cryptography, which has developed in recent years, or understanding the origin of data with differential privacy.

The combination of different data processing techniques will lead to a lowering of the ϵ parameter by making it possible to use the data and making this type of process more attractive to more sceptical companies.

If this type of privacy becomes more and more popular, users around the world will have a greater sense of security and will be more inclined to release their data, including the most sensitive aspects, thus leading new statistical research to benefit from this valuable advantage.

Appendix

An example of the Rstudio code used to carry out the analysis in Chapter 4 is presented below, in particular, only $\epsilon = 3$ and $n = 100$ are treated.
The other parameter values and dataset size are replaced in the same procedure.

```
library(diffpriv)
#1) Laplace mechanism
#a) Dataset1
n<-100
n1<-10
n2<-10
D <- matrix(runif(n, min = 0, max = 1),nrow=n1,ncol=n2) #database [0,1]^n
f <- function(X) mean(X) # target function --> mean
mechanism <- DPMechLaplace(target = f, sensitivity = 1/n, dims = 1)
pparams <- DPParamsEps(epsilon = 3) ## desired privacy budget
r <- releaseResponse(mechanism, privacyParams = pparams, X = D)
cat("Private response :", r$response,
    "Non-private response f(D): ", f(D2))
error<- abs(f(D)-r$response)
#2) Gaussian mechanism
#a) Dataset1
N<-100
N1<-10
N2<-10
```

```
G1 <- matrix(runif(N, min = -0.5, max = 0.5),nrow=N1,ncol=N2) #database [-1/2,1/2]^n
f1<- function(X) mean(X)
mechanism1 <- DPMechGaussian(target = f1, sensitivity = 1/N, dims = 1)
ppara <- DPParamsDel(epsilon = 1, delta = 10^(-4)) ## desired privacy budget
r1 <- releaseResponse(mechanism1, privacyParams = ppara, X = G1)
d1<-r1$response
cat("Private response :", r1$response,
    "Non-private response f(D): ", f1(G1))
#3) Bernstein Mechanism
#Dataset 1) N=100, epsilon 3,5
pck_regression <- function(D, bandwidth = 0.1) {
    K <- function(x) exp(-x^2/2)
    ids <- sort(D[,1], decreasing = FALSE, index.return = TRUE)$ix
    D <- D[ids, ]
    n <- nrow(D)
    ws <- (D[2:n,1] - D[1:(n-1),1]) * D[2:n,2]
    predictor <- function(x){
    sum(ws * sapply((x - D[2:n,1]) / bandwidth, K)) / bandwidth
    }
return(predictor)
}

N <- 100
D <- runif(N)
D <- cbind(D, sin(D*10)*D + rnorm(N, mean=0, sd=0.2))
model <- pck_regression(D)
K <- 25
bmodel <- bernstein(model, dims=1, k=K)
m <- DPMechBernstein(target=pck_regression, latticeK=K, dims=1)
P <- function(n) { # a sampler of random, "plausible", datasets
```

```
    Dx <- runif(n)
    Dy <- rep(0, n)
    if (runif(1) < 0.95) Dy <- Dy + Dx
    if (runif(1) < 0.5) Dy <- Dy * sin(Dx)
    if (runif(1) < 0.5) Dy <- Dy * cos(Dx)
    cbind(Dx, Dy + rnorm(n, mean=0, sd=0.2))
}
m <- sensitivitySampler(m, oracle=P, n=N, gamma=0.20, m=500)
R <- releaseResponse(m, privacyParams=DPParamsEps(epsilon=3), X=D)
R1<- releaseResponse(m, privacyParams=DPParamsEps(epsilon=1), X=D)
pmodel <- R$response
pmodel1<- R1$response
xs <- seq(from=0, to=1, length=50)
yhats <- sapply(xs, model)
yhats.b <- predict(bmodel, xs)
yhats.p <- R$response(xs)
y1hats.p<- R1$response(xs)
```

Acknowledgements

At the conclusion of this work I feel it is my duty to thank all the people who have contributed and supported me along this university career.

First of all, I would like to thank the supervisor of this thesis, Professor Formentin, for his availability, patience, support and help during the writing of this thesis.

I would like to thank my parents and my sister Sara, for allowing me to achieve this degree, supporting me at all times and never ceasing to believe in me, giving me the strength to face this path.

I would also like to thank Marco, as well as my greatest supporter and the one who gave me the strength, constancy and perseverance to face this path without never giving up, always believing in me.

My thanks also go to my family and all the people who have been close to me, with whom I have shared every moment of joy, but also of discouragement during these five years together.

List of Figures

List of Tables

Bibliography

[1] Dwork, Cynthia, and Aaron Roth. "The algorithmic foundations of differential privacy." Foundations and Trends® in Theoretical Computer Science 9.3–4 (2014): 211-407.

[2] John Duchi, Lecture Notes, 2019.

[3] Europea, Unione. "Regolamento UE 2016/679 del Parlamento europeo e del Consiglio del 27 aprile 2016 relativo alla protezione delle persone fisiche con riguardo al trattamento dei dati personali, nonché alla libera circolazione di tali dati e che abroga la direttiva 95/46." CE (regolamento generale sulla protezione dei dati).

[4] Achille Pierre Paliotta, La de-anonimizzazione dei dati personali. Il caso del dataset Netflix, 2021. https://www.ictsecuritymagazine.com/articoli/la-de-anonimizzazione-dei-dati-personali-il-caso-del-dataset-netflix/

[5] Hossein Pishro-Nik, Introduction to Probability, Statistics, and Random Processes, 2014.

[6] Kaissis, Georgios, et al. "A unified interpretation of the gaussian mechanism for differential privacy through the sensitivity index." arXiv preprint arXiv:2109.10528 (2021).

[7] Gautam Kamath, Lecture 4, Intro to Differential Privacy, Part 2, Algorithms for Private Data Analysis, 2020.

[8] Benjamin Steephenson, Implementing Differential Privacy Using Randomized Response Algorithms, Tufts University, Introduction to Computer Security, 2017. https://www.cs.tufts.edu/comp/116/archive/fall2017/bsteephenson.pdf

[9] Mironov, Ilya. "Rényi differential privacy." 2017 IEEE 30th computer security foundations symposium (CSF). IEEE, 2017.

[10] Narayanan, Arvind, and Vitaly Shmatikov. "Robust de-anonymization of large sparse datasets." 2008 IEEE Symposium on Security and Privacy (sp 2008). IEEE, 2008.

[11] Slavkovic, Aleksandra, and Fei Yu. "O privacy, where art thou?: genomics and privacy." Chance 28.2 (2015): 37-39.

[12] Wang, Shuang, et al. "Genome privacy: challenges, technical approaches to mitigate risk, and ethical considerations in the United States." Annals of the New York Academy of Sciences 1387.1 (2017): 73-83.

[13] Microsoft, What is Differential Privacy in Machine Learning?, 2022. https://www.microsoft.com/en-us/ai/ai-lab-differential-privacy

[14] Andrew Hutchinson, Facebook Outlines New Differential Privacy Framework to Protect User Information in Shared Datasets, 2020. https://www.socialmediatoday.com/news/facebook-outlines-new-differential-privacy-framework-to-protect-user-inform/579167/.

[15] Girgis, Antonious M., et al. "On the renyi differential privacy of the shuffle model." Proceedings of the 2021 ACM SIGSAC Conference on Computer and Communications Security. 2021.

[16] Dwork, Cynthia, Guy N. Rothblum, and Salil Vadhan. "Boosting and differential privacy." 2010 IEEE 51st Annual Symposium on Foundations of Computer Science. IEEE, 2010.

[17] McSherry, Frank D. "Privacy integrated queries: an extensible platform for privacy-preserving data analysis." Proceedings of the 2009 ACM SIGMOD International Conference on Management of data. 2009.

[18] Dwork, Cynthia, et al. "Calibrating noise to sensitivity in private data analysis." Theory of cryptography conference. Springer, Berlin, Heidelberg, 2006.

[19] Rubinstein, Benjamin IP, and A. Francesco. "diffpriv: An R package for easy differential privacy." Journal of Machine Learning Research 18 (2017): 1-5.

[20] Alex Brasch, A Comparison of Census 2010 SF1 Differentially Private Data in Oregon, Winter Quarter, 2020.

[21] Alda, Francesco, and Benjamin IP Rubinstein. "The bernstein mechanism: Function release under differential privacy." Thirty-First AAAI Conference on Artificial Intelligence. 2017.

[22] Balle, Borja, and Yu-Xiang Wang. "Improving the gaussian mechanism for differential privacy: Analytical calibration and optimal denoising." International Conference on Machine Learning. PMLR, 2018.

[23] Joy, Kenneth I. "Bernstein polynomials." On-Line Geometric Modeling Notes 13.4 (2000).

[24] Rubinstein, Benjamin IP, and Francesco Aldà. "Pain-free random differential privacy with sensitivity sampling." International Conference on Machine Learning. PMLR, 2017.

[25] Herawati, Netti, et al. "The Nonparametric Kernel Method using NadarayaWatson, Priestley-Chao and Gasser-Muller Estimators for the Estimation of the Rainfall Data in Lampung." International Journal of Mathematics Trends and Technology 68.8 (2022): 12-20.