Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
CHAPTER 3

COVID-19’s pandemic: a new way of thinking through linear combinations of proportions

Adelaide Freitas\(^a\), Sara Escudeiro\(^b\), Juliana Reis\(^a\), and Cristiana J. Silva\(^a\)

\(^a\)Center for Research and Development in Mathematics and Applications (CIDMA), Department of Mathematics, University of Aveiro, Aveiro, Portugal \(^b\)College of Agriculture (ESAC), Polytechnic Institute of Coimbra, Coimbra, Portugal

3.1 Introduction

The Coronavirus disease 2019, most commonly known as COVID-19, is a severe acute respiratory syndrome caused by the SARS-CoV-2 virus. The first COVID-19 outbreak was reported in China in December 2019 and then the disease quickly spread across the world, causing a large number of deaths worldwide and thereby posing a serious threat to public health [13,15]. Since December 31, 2019, until the end of the first week of September 2021, approximately 225 million cases of COVID-19 have been reported, including more than 4.6 millions of deaths, worldwide [7]. In Portugal, the first two cases confirmed with COVID-19 disease were reported on March 2, 2020, while the first death was documented on March 16, 2020 [5]. On March 18, 2020, the first state of emergency was declared in Portugal, demanding for imposing the application of extraordinary and urgent restrictive measures regarding movement rights and economic freedoms, with the purpose of preventing the transmission of the virus [1]. This state of emergency lasted 45 days and ended on May 2. After the restrictive measures applied countrywide between 18 March and 2 May, which included lockdowns, social distancing and mask mandates, the epidemic crisis waned and COVID-19 disease burden reached its “lowest level” during the summer period of 2020. In this so-called lower level of COVID-19 disease burden, there were about 200–300 new daily cases, most of which were mild and didn’t require hospitalization, with less than 1% of the active cases hospitalized in Intensive Care Units (ICU) and about 0–13 deaths per day. During the summer period of 2020 until the next new state of emergency (November 6, 2020), the Portuguese government alternately implemented states of contingency or states of alert, depending on the countrywide or
regional epidemic situation [2]. In September of 2020, a steady increase of the infected people marked the beginning of a second wave that peaked in November, leading to a “critical level” of COVID-19 disease burden characterized by more than 6000 daily new cases and a shortage of beds in ICU of some public hospitals, in spite of the reintroduction of some restrictive measures in late October and the beginning of November of 2020 [3]. By September 17, 2021, the Directorate-General of Health (DGS) of Portugal, has documented a total of 1,060,432 confirmed cases of COVID-19 and 17,895 deaths by COVID-19 [5]. In this paper, we focus on COVID-19 real data of Portugal, for the period between March 2, 2020 and April 16, 2021, but it may be applied to other real data of a different country or communicable disease.

COVID-19 is a new disease with potentially serious effects on the health of individuals, many of which are still unknown to the scientific community. Consequently, the timely definition of action plans to prevent and fight the disease is urgently needed. It is therefore important that policy makers have the necessary tools to better forecast the capability of the National Health System (NHS) to response to the spread of the pandemic virus and the mobilization of human resources to strengthen the virus tracking capacity. In this sense, knowledge of the evolution of the number of nonhospitalized and hospitalized patients (in infirmary and ICU) and the occurrence of extreme situations (deaths resulting from the disease) are critical.

SARS-CoV-2 spreading has been the central theme of vast recent research works. In the area of mathematics, it is quite common to find references to deterministic compartmental models, such as the SIR-type models, in attempts to describe the transmission dynamics of the virus and the evolution of this epidemic outbreak in specific regions of the world, trying to predict the appearance of new waves and the number of reported and unreported cases, among other indicators [12,14,18].

In this work, we propose a new approach to analyze the trajectory of epidemic curves based on the behavior over time of estimated adequate linear combinations of proportions related to (hospitalized or nonhospitalized) active infected individuals with COVID-19 and deaths due to this disease per day. Assuming that these daily proportions are parameters from populations which can be modeled by independent binomial distributions, confidence intervals of these linear combinations are constructed. These interval estimates provide information about the daily evolution of the experienced epidemiological situation and, consequently, enable a evaluation of the impact of more or less restrictive measures that can be applied as prevention and control strategies. To the best of our knowledge, this is the first time that linear combinations of proportions have been applied to the study of the spreading of virus of COVID-19 disease and its consequences on the public health systems. We focus our analysis using daily epidemiological data from COVID-19 officially made available by DGS of Portugal [5].
3.2 Estimation of linear combinations of proportions

Inference for one binomial proportion \( p \) is one of the most basic analyses in Statistics. There are several methods for the construction of confidence intervals (CI) for one proportion \( p \), being the score method (also called Wilson method) one of the most known [10,20]. Due to its important practical value, extensions for the difference of two independent binomial proportions and, although in smaller number, for any linear function of \( k > 2 \) success proportions of independent binomial populations have received some attention. In particular, new CI methods for linear combinations of proportions have been emerged by considering different types of adjustment of classic methods for estimating single binomial proportions [16,17]. For instance, [19] proposed a hybrid score CI for the difference between two proportions, \( p_1 - p_2 \), taking into account the limits of the CIs for each of the proportions \( p_1 \) and \( p_2 \) given when the score method is applied. Later, [22] named this approach as Method Of Variance Estimates Recovery (MOVER) and extended it for any linear combination of \( k \geq 1 \) independent binomial proportions. The MOVER approach provides CI with coverage probabilities close to nominal level [11,22].

Formally, let \( X_1, \ldots, X_k \) be \( k \geq 2 \) independent binomial random variables with parameters \( n_i \) and (unknown) population proportion \( p_i \), with \( i = 1, 2, \ldots, k \). Considering a linear combination of binomial proportions defined as \( L = \sum_{i=1}^{k} \beta_i p_i \), where \( \beta_i \neq 0 \) is a fixed value, the maximum likelihood estimate (MLE) of \( L \) is given by \( \hat{L} = \sum_{i=1}^{k} \beta_i \frac{X_i}{n_i} \) and the variance of \( \hat{L} \) is given by \( v(\hat{L}) = \sum_{i=1}^{k} \beta_i^2 p_i (1 - p_i) n_i \). In a general framework, the bounds of the score CI for any parameter \( \theta \) are determined by solving the following equation in order to \( \theta \):

\[
\frac{\hat{\theta} - \theta}{\sqrt{v(\hat{\theta})}} = z_{\alpha/2},
\]

where \( z_{\alpha/2} \) is the \( \alpha/2 \) upper quantile of the standard normal distribution, \( \hat{\theta} \) is the MLE of \( \theta \) and \( v(\hat{\theta}) \) represents the variance of the estimator \( \hat{\theta} \). Consequently, to compute the lower and upper bounds of the CI for one proportion we need to solve Eq. (3.1) when \( \theta \) is one proportion (parameter \( p_i \)). The MLE of \( p_i \) is \( \hat{p}_i = \frac{X_i}{n_i} \) and its variance is given by \( v(\hat{p}_i) = \frac{\hat{p}_i (1 - \hat{p}_i)}{n_i} \). Therefore, after algebraic calculations, we get the lower \( (l_{i-}) \) and upper \( (l_{i+}) \) bounds of the Wilson CI [21] for the single proportion \( p_i \) that are given by:

\[
l_{i\pm} = \frac{\hat{p}_i n_i + \frac{z_{\alpha/2}^2}{2} \mp \frac{z_{\alpha/2}^2}{4} + \hat{p}_i (1 - \hat{p}_i) n_i}{n_i + \frac{z_{\alpha/2}^2}{4}}.
\]
Thus, taking into account these limits of the CIs for each of the proportions \( p_i, i = 1, 2, \ldots, k \), the lower and upper limits of the MOVER CI for \( L = \sum_{i=1}^{k} \beta_i p_i \), are given by [22]:

\[
\hat{L} \pm z_{\alpha/2} \sqrt{\sum_{w_i < 0} \frac{w_i^2 l_i (1 - l_i)}{n_i} + \sum_{w_i > 0} \frac{w_i^2 l_i (1 - l_i)}{n_i}}.
\]

(3.3)

In this work, formula (3.3) will be used to obtain interval estimates of interpretable linear combinations of proportions of daily events related to the evolution of the COVID-19 pandemic in Portugal, for the period between March 2, 2020 and April 16, 2021.

3.3 Material and methods

3.3.1 Datasets

The data used in this work was extracted from different data sources. One of them contains daily count data related to the number of infected persons that are at home or hospitalized in an infirmary or ICU, and the daily number of deaths by COVID-19, from March 2, 2020 to April 16, 2021. This dataset was provided by Data Science for Social Good Portugal and DGS of Portugal [5,8]. The dataset containing the total number of deaths per day during the period under study was obtained from [9].

The number of infirmary and ICU beds specifically assigned to COVID-19 patients was retrieved from several governmental and news websites, as this information was not published in a structured fashion by the relevant official entities. The number of beds for this purpose was increased several times by the NHS during the epidemic to keep up with the increasing number of infected people requiring hospitalization: March to September 2020 – 1500 infirmary beds, 250 ICU beds; October 2020: 2047 infirmary beds, 356 ICU beds; November 2020 to April 2021 – 2047 infirmary beds, 373 ICU beds. It is worth to mention that these numbers were made available through the media (e.g., [4,6]), and so, they may not be exact.

3.3.2 Linear functions of proportions

In order to study the evolution of the pandemic situation in Portugal during the period under analysis, linear combinations considered relevant of the following proportions were constructed:

- \( p_1(t) \), the proportion of COVID-19 nonhospitalized active infected individuals, at day \( t \), given by

\[
p_1(t) = \frac{\text{number of nonhospitalized active infected individual}}{\text{total number of active infected individuals}};
\]
COVID-19’s pandemic: a new way of thinking through linear combinations of proportions

- \( p_2(t) \), the proportion of COVID-19-associated hospitalization, at day \( t \), given by
  \[
  p_2(t) = \frac{\text{number of COVID-19 hospitalized active infected individuals}}{\text{total number of allocated hospital beds for COVID-19}}
  \]

- \( p_3(t) \), the proportion of death by COVID-19, at day \( t \), given by
  \[
  p_3(t) = \frac{\text{number of individuals which die by COVID-19 disease}}{\text{total number of deaths}}
  \]

- \( p_4(t) \), the proportion of COVID-19 related infirmary occupancy, at day \( t \), given by
  \[
  p_4(t) = \frac{\text{number of COVID-19 infirmary beds for COVID-19 currently occupied}}{\text{total number of infirmary beds reserved for COVID-19}}
  \]

- \( p_5(t) \), the proportion of COVID-19 related ICU occupancy, at day \( t \), given by
  \[
  p_5(t) = \frac{\text{number of COVID-19 ICU beds for COVID-19 currently occupied}}{\text{total number of ICU beds reserved for COVID-19}}
  \]

Note that, for the calculation of the proportion referring to the number of deaths by COVID-19, the total number of deaths per day was used in the denominator, in order to take into account the real impact of COVID-19 deaths on the total number of daily deaths.

The behavior of interpretable linear combinations of these proportions over time are then investigated. To unify the notation, a general linear function \( L \), given by

\[
L(t; \omega) = \sum_{i=1}^{5} \omega_i p_i(t),
\]

is constructed. Obviously, depending on the weight vector \( \omega = (\omega_1, \ldots, \omega_5) \), many different linear combinations of the proportions \( p_i(t), i = 1, \ldots, 5 \), can be established. For some particular choices of \( \omega \), different linear functions \( L \) could be considered providing useful information. The simplest ones are produced when there is only a single nonnull weight \( \omega_i \), which is equal to the unit. For instance, \( L(t; 1, 0, 0, 0, 0) = p_1(t) \) corresponds to the proportion of nonhospitalized active cases. This type of cases does not represent a burden to the NHS. When two or three components in the weight vector \( \omega \) are nonnull, examples of interpretable linear combinations of two or more proportions can also be described. For the case of two proportions, some examples are listed in Table 3.1. For three proportions, two particular settings of weight vectors \( \omega \), and consequently linear functions \( L \), will be herein considered, namely:
Table 3.1: Interpretation of $L(t; \omega)$ for particular cases of $\omega = (\omega_1, \ldots, \omega_5)$ with only two nonnull weights $\omega_i$.

| $\omega_i \neq 0$ | Interpretation of $L(t; \omega) = \omega_1 p_1(t) + \omega_2 p_2(t) + \omega_3 p_3(t) + \omega_4 p_4(t) + \omega_5 p_5(t)$ |
|-------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| for $i = 1, 2$    | Related to all active infected individuals. For instance, $L(t) = p_1(t) - p_2(t)$ is aimed for comparing the proportions of hospitalized and nonhospitalized active infected individuals. Negative values for $L$ means an unfavorable situation for the NHS since the hospital occupancy rate by COVID-19 patients exceeds the percentage of nonhospitalized cases. |
| for $i = 2, 3$    | For assessing hospitalization and deaths. For instance, $L(t) = \frac{1}{2}(p_2(t) + p_3(t))$ reflects to some extent the average capability of the NHS to respond to the epidemic crisis. An increase in the value of $L$ reflects an increased burden on the NHS. Furthermore, $L(t) = p_2(t) - p_3(t)$ allows for comparing the proportions of hospitalized active infected individuals and death due to COVID-19 disease. Negative values for $L$ means an unfavorable scenario since death by COVID-19 is more likely than hospital beds being occupied with infected patients. |
| for $i = 4, 5$    | For comparing the response capacity of the NHS between infirmary and ICU. For instance, $L(t) = p_4(t) - p_5(t)$ allows for comparing the proportions of hospitalized active infected individuals that are at infirmary and ICU. Negative values for $L$ means an unfavorable scenario. The lower this value, the more worrying is the hospital situation experienced. An active infected individual admitted to infirmary is not in as weak health situations as those admitted to the ICU. |
| for $i = 3, 5$    | For assessing the most critical factors (hospitalization at ICU or death). For instance, $L(t) = \frac{1}{2}(p_3(t) + p_5(t))$ reflects to some extent the capability of the NHS to respond to the two worst consequences of the pandemic crisis. An increase in the value of $L$ reflects an increased burden on the NHS. |

- $L(t; \omega_1, \omega_2, \omega_3, 0, 0) = \omega_1 p_1(t) + \frac{1}{2}(\omega_2 p_2(t) + \omega_3 p_3(t))$;
  Interpretation: It allows the comparison of the evolution of the proportion of infected persons in domestic environment versus a hospital scenario or death. If the value of the proportion of nonhospitalized patients is much higher than the average of the proportions of hospitalized patients and deaths, we can say that Portugal is in a favorable situation regarding the pandemic. The opposite situation occurs if the average of the proportions of hospitalized patients and deaths is greater than an adequate weighted proportion of nonhospitalized patients.

- $L(t; 0, 0, \omega_3, \omega_4, \omega_5) = \omega_3 p_3(t) + \frac{1}{2}(\omega_4 p_4(t) + \omega_5 p_5(t))$;
  Interpretation: It allows the comparison of the evolution of the proportion of active infected persons admitted to infirmary and ICU versus the death scenario. This means that this function $L$ describes the hospital situation during the epidemic episode against the number of deaths by COVID-19.

Concretely, in the present study, the following types of linear combinations are analyzed:
COVID-19’s pandemic: a new way of thinking through linear combinations of proportions 47

Table 3.2: Interpretation of equalities resultant from $L_j(t) = 0$, $j = 1, \ldots, 5$.

| Equation                                      | Interpretation for COVID-19 patients                                      |
|-----------------------------------------------|-------------------------------------------------------------------------|
| $p_3(t)^2 = \omega_1 p_1(t)$                  | The proportion of deaths is a fraction $\omega_1$ of the proportion of hospitalized patients. |
| $p_4(t)^2 = \omega_2 p_2(t)$                  | The proportion of patients in ICU is a fraction $\omega_2$ of those in infirmary. |
| $p_5(t)^2 = \omega_3 p_3(t)$                  | The average of the proportions of patients in critical situation (dead or in ICU) is a fraction $\omega_3$ of the proportion of sick in infirmary. |

- $L_1(t) \equiv L(t; \omega_1, -1/2, -1/2, 0, 0) = \omega_1 p_1(t) - \frac{p_3(t)^2 + p_5(t)^2}{2}$;
- $L_2(t) \equiv L(t; \omega_1, -1, 0, 0, 0) = \omega_1 p_1(t) - p_2(t)$;
- $L_3(t) \equiv L(t; 0, \omega_2, -1, 0, 0) = \omega_2 p_2(t) - p_3(t)$;
- $L_4(t) \equiv L(t; 0, 0, 0, \omega_4, -1) = \omega_4 p_3(t) - p_5(t)$;
- $L_5(t) \equiv L(t; 0, 0, -1/2, \omega_4, -1/2) = \omega_4 p_4(t) - \frac{p_3(t)^2 + p_5(t)^2}{2}$.

In each combination, one weight $\omega_i$ was kept free. In terms of interpretability, these combinations are aimed at assessing domestic and hospital (infirmary, ICU or both) scenarios or deaths related to COVID-19 disease, at the time $t$. The free weight in each combination has the advantage of providing flexibility for comparing probabilistic relations between two COVID-19 related events of interest. Making each of the above combinations equal to zero, interpretable relationships can emerge (Table 3.2).

3.3.3 Inference

Although the COVID-19-based datasets have been provided by official sources, these same sources frequently publish revisions to those datasets. In order to take into account some possible lack of accuracy of the data (due to measurement errors and/or bias), in addition to calculating estimates of each of the five linear combinations $L_1(t), \ldots, L_5(t)$, for each day $t$, their corresponding CI were also constructed, using the MOVER approach (formula (3.3)). Indeed, the sample is not random and is skewed, as the testing for COVID-19 is made in specific groups of the Portuguese population. For achieving a golden data status, all Portuguese population should have been tested. As we will most probably not have early access to such kind of (complete) data, our studies will be based on official available datasets (see Section 3.2), for which we assume they are representative of the pandemic situation in Portugal. We also assume the $L_j$-values observed for each $t$ (unit of time used: day) is a realization of a random variable that follows a binomial distribution, as theoretically described in Section 3.2.
3.3.4 Graphical procedure

The five linear functions, $L_1, \ldots, L_5$, were graphically displayed in order to allow the analysis of the behavior of the epidemic over time. Two different plots were considered for each function.

Firstly, trajectories of the function $L_j(t)$, over $t$, for different values of the free weight were plotted. Graphically, it corresponds to vertically shifting the trajectory of the series. Due to the interpretation of the case $L_j(t; \omega) = 0$, we considered the value of the free weight of $L_j$ for which the MOVER-based CI at 95% level of confidence for $L_j(t)$ includes the zero value in most days $t$ between the middle of June and early September 2020, approximately the summer period of 2020 referenced by a lower level of COVID-19 disease burden occurred in Portugal (reference period for a favorable situation). Hence, this weight was determined experimentally. Furthermore, for better framing each period of the pandemic, the following important dates were highlighted in these trajectory graphs using vertical lines: 18 Mar, 2020 – First state of emergency; 2 May, 2020 – First state of emergency canceled; 14 Oct, 2020 – State of calamity; 9 Nov, 2020 – New state of emergency; 15 Jan, 2021 – New lockdown period; 15 Mar, 2021 – Easing of some lockdown restrictions.

Afterwards, a scatter-plot of the points $(x, y)$ was plotted for each function $L_j$, where $x$ ($y$) represents the first (second, resp.) member of the equation indicated in Table 3.2, equivalent to $L_j(t) = 0$ and with unit free weight. These points were chronologically joined by segments using a gradient color scale in order to visually explore the trajectory over $t$ of proportions of events considered on $L_j(t)$. Additionally, the scatter-plots allow analyzing each marginal per si.

3.4 Results and discussion

3.4.1 Linear combination $L_1(t) = \omega_1 p_1(t) - \frac{p_2(t) + p_3(t)}{2}$

The plot of trajectories of the linear function $L_1$ visually facilitates the comparison between a comfortable scenario for the NHS, in terms of not using hospital resources (patients are infected, but at home, $p_1(t)$) versus an unfavorable scenario (patients are infected and hospitalized, $p_2(t)$, or dead, $p_3(t)$) related to COVID-19 pandemic situation. The level of resource utilization of the NHS was based on the total number of beds reserved for COVID-19 patients occupied, that is, making no distinction between ICU and infirmary beds.

Plots of $L_1(t)$ and its 95%-CI, for the period from March 2, 2020 to April 16, 2021, led us to experimentally set the weight to $w_1 = 0.125$ as this was a value that visually adjusted the trajectory of $L_1$ to the reference value of zero during the summer period. During this period,
the hospital resources utilization was considered as comfortably under control by the NHS. Technically, as mentioned, this value of $\omega_1$ was determined such that the zero value of $L_1(t)$ belongs to (or surpass higher) the CI in most days $t$ of the summer period of 2020 of the favorable situation for the NHS (Fig. 3.1). Therefore, a “favorable situation” occurs when

$$\frac{p_2(t) + p_3(t)}{2} \approx 0.125 p_1(t). \tag{3.4}$$

Relation (3.4) establishes that the pandemic scenario can be described as “comfortably sustainable” for the NHS whenever the average of the proportions of the worst conditions (hospitalized and dead) is approximately equal to (or does not exceed) about 12.5% of the percentage of infected patients at home.

After Portugal reached the beginning of the summer period (favorable situation), the Portuguese government started to ease the restrictive measures that it has put in place when the epidemic first affected the country. After the end of the summer, it is possible to observe, from the trajectory in Fig. 3.1, a deterioration of the epidemic situation, which reached a deeper level in the period from November to January and, subsequently, an even deeper peak in the beginning of February, which coincides with the most negative $L_1(t)$ values obtained (the most critical unfavorable situation for the NHS during the pandemic crisis). From March 2021 on, the situation became increasingly favorable evolving into a situation similar to that of the summer 2020 period (favorable situation).

Now, considering the period between November 2020 to January 2021, which was long enough to be used as a reference period and corresponds to an unfavorable scenario for the
NHS (system responsiveness at its limit or outside it), the weight $w_1 = 0.6$ was also chosen based on experimentation according to the same reasoning as before, that is, to make the 95%-CI of $L_1(t)$ include the zero value in most days $t$ during that unfavorable period. Graphically, it corresponds to a vertical shifting of the time series in Fig. 3.1. Therefore, during that unfavorable situation,

$$p_2(t) + p_3(t) \approx 0.6p_1(t),$$

which means that the average of the proportions of the worst conditions (hospitalized and dead) represents, approximately, 60% of the proportion of infected persons at home, an increase of about 50% compared to the favorable scenario described in (3.4), thus showing the need for governmental measures to prevent the unfavorable progression of the pandemic.

In the scatter-plot of Fig. 3.2, there is a line corresponding to the trajectory of the pair formed by $p_1$ and $p_2 + p_3$ for the period between March 2, 2020 and April 16, 2021 and two straight lines that represent, respectively, the boundary between the favorable ($w_1 = 0.125$) and unfavorable ($w_1 = 0.6$) situations, as mentioned above. The slopes of these two straight lines are
equal to these weights. Points near, but below, the straight line of slope 0.125 correspond to favorable period for the NHS. Analyzing per marginal, it is clear that, from the end of April 2020, almost all infected active persons are at home \( p_1(t) > 0.95 \) and during the unfavorable situation (November 2020 to January 2021) the mean of the two rates, the hospital occupancy rate by COVID-19 patients and the death rate due to COVID-19 disease, was higher than 50%, having decreased after the lockdown measures imposed in January 2021. A value around 20% was achieved for this mean in the middle of March 2021.

### 3.4.2 Linear combination \( L_2(t) = \omega_1 p_1(t) - p_2(t) \)

In terms of hospital resource utilization, it is more convenient for the NHS that a COVID-19 infected person stays at home. Taking the summer period as the reference for what is considered a favorable situation in the country, the following relation was estimated from Fig. 3.3: \( p_2(t) \approx 0.25 p_1(t) \). It means that if we consider only the hospital beds reserved for COVID-19 infected persons, the NHS can comfortably deal with the challenges raised by the COVID-19 pandemic if the occupancy rate represents approximately (or does not exceed) about 25% of the proportion of active cases not requiring hospitalization. Clarifying a little further, since \( p_1(t) \leq 1, \forall t \), an occupancy rate not exceeding 25% of the total number of beds reserved for COVID-19 patients turned out to be comfortable for the NHS.

Continuing the analysis of Fig. 3.3, we can see that in October 2020, the hospital occupancy rate by COVID-19 patients reached a value of 1/4 of the percentage of nonhospitalized active cases, marking the beginning of the trend that led to a new unfavorable situation, that is, an increase in the active hospitalized cases. The NHS reacted by further increasing the number of beds reserved for COVID-19 patients from the beginning of November and hiring additional
medical personnel to staff them. This trend, highly pronounced in January, reversed its direction 15 days after a new lockdown was imposed on January 15th, 2021.

Fig. 3.4 allows to complement the analysis of Fig. 3.3. In Fig. 3.4, it is possible to observe that the hospital occupancy rate of the beds reserved for COVID-19 patients reaches its upper limit during the unfavorable period between November 2020 and January 2021, decreasing to 30% in the middle of March 2021 when some of the restrictions imposed during the lockdown decreed by the Portuguese government in January 15th was lifted. Furthermore, Fig. 3.4 presents a shape similar to Fig. 3.2, with a wider range of values in the y-axis. This comparison allows us to deduce that the proportion of deaths by COVID-19 ($p_3(t)$) is much lower than the proportion of hospitalized cases ($p_2(t)$), thus decreasing the value of their average for each $t$. However, since the graphs in Figs. 3.1 and 3.3 have also similar shape but function $L_1$ exhibits a slower rate of change from around the beginning of February ($\pm 15$ days), a great increase in the proportion of deaths by COVID-19 ($p_3(t)$) during that specific period is deduced. Indeed, this can be clearly confirmed by the analysis of the behavior of the marginal $p_3(t)$ in Fig. 3.6, where it is possible to observe that the value of this proportion is generally around 0.25 and reaches its maximum value of approximately 0.47 after 15 days, on January 31st, 2021.

### 3.4.3 Linear combination $L_3(t) = \omega_2 p_2(t) - p_3(t)$

The linear function $L_3$ deals with the two most critical types of events of the pandemic crisis: hospitalizations and deaths by COVID-19. For $\omega_2 = 0.07$ and $\omega_2 = 0.20$, the corresponding function $L_3(t)$ remained approximately constant around zero during the summer period (Fig. 3.5) and the month of November (plot not shown), respectively. Hence, during the “comfortable” and the “uncomfortable” periods for the NHS, the proportion of deaths by COVID-
Figure 3.5: Selection of $\omega_2 = 0.07$ to get a trajectory of the linear combination $L_3(t)$ with values close to zero in the summer period.

Figure 3.6: Trajectory of the pair $(p_2, p_3)$ during the period under study.

COVID-19 showed to be equal to 7 and 20 percent, respectively, of the hospital occupancy rate by infected persons. In Fig. 3.5, when $L_3$ becomes negative, a trend that started a few days before the declaration of the state of emergency, on October 14th, it indicates an increase in mortality by COVID-19. The proportion of deaths $p_3(t)$ reaches a value of 7% of the COVID-19 associated hospital occupancy rate in the beginning of April, two weeks after some of the lockdown restrictions were lifted.

From the analysis of Fig. 3.6, it is possible to observe that the death rate by COVID-19 increased by about 15% on November 9th, when a new state of emergency was declared. Following the restrictive measures introduced by the state of emergency, the death rate by COVID-19 stabilized until the beginning of 2021, when it began to deteriorate substantially (most probably due to negligent conducts during the Christmas and New Year celebrations), despite a new confinement having been decreed on January 15th, and ended up reaching approximately 47% on January 31th, as already mentioned above. Having reached this peak, the
death rate began to decline steadily from February 18th, remaining below 20% thereafter. Despite this, the hospital occupancy rate associated with COVID-19 remained close to its upper limit. This reminds us the importance of increasing the number of hospital beds, especially ICU beds, as well as the number of professionals specialized in intensive care required to staff those beds, to overcome the shortages reported by the media.

### 3.4.4 Linear combination $L_4(t) = \omega_4 p_4(t) - p_5(t)$

Taking as reference the summer period, the weight $\omega_4 = 1$ shows to be adequate. Setting this weight, the linear combination $L_4(t)$ represents the difference of two proportions that makes it possible to analyze the hospital situation in relation to hospitalized patients infected with COVID-19, being discriminated against patients hospitalized in the infirmary and patients hospitalized in ICU. The results obtained are shown in Figs. 3.7 and 3.8. Patients with more severe conditions are admitted to an ICU and, as such, this kind of admission is considered to be an unfavorable situation.

From Fig. 3.7, it is possible to assume that during the first wave (around April 20th, 2020), having $L_4(t) < 0$ means that the utilization rate of the COVID-19 dedicated ICU overall capacity during the pandemic is greater than the same rate for infirmary capacity. The null value of $L_4(t)$ for the period between November 9th, 2020 and the middle of February, 2021 is well described in Fig. 3.8, which shows $p_4(t) \approx p_5(t)$. By analyzing Fig. 3.8, we can conclude that the NHS exceeded its capacity in terms of both infirmary and ICU beds dedicated to COVID-19 patients. Still in Fig. 3.8, having points $(x, y)$ above the straight line $y = x$, which corresponds to $L_4(t) = 0$ (the summer period), indicates the existence of periods where the occupancy rate of ICU beds was higher than that of infirmary, taking into account each hospital’s capacity.
Figure 3.8: Trajectory of the pair \((p_4, p_5)\) during the period under study.

### 3.4.5 Linear combination

\[ L_5(t) = \omega_4 p_4(t) - \frac{p_3(t) + p_5(t)}{2} \]

This linear function relates the proportion of infected people admitted to infirmary, the proportion of infected people admitted to ICU and the proportion of deaths by COVID-19. Among these three events, admitting patients to infirmary is considered to be the favorable situation. Experimentally, the value \(\omega_4 = 0.5\) was found to be an adequate choice for achieving the relation \(L_5(t) \approx 0\) during the summer period, allowing us to conclude that

\[ p_5(t) \approx p_4(t) - p_3(t), \tag{3.5} \]

for this period. Therefore, considering the occupancy rate of hospital resources, relation (3.5) induces that the ICU \((p_5(t))\) should be lower than that for the infirmary for a “comfortable” hospital pandemic scenario.

By analyzing the results presented in Fig. 3.9, it is observed that at the beginning of both first and second waves, in April, 2020 and in November, 2020, respectively, there was a very sharp drop in the \(L_5\) values, which translates into a deterioration of the hospital situation. After the measures applied by the government on March 18th, 2020 to May 2nd, 2020 \(L_5\) values started increasing and then stabilizing until September, 2020, when they again started to fall. Curiously, \(L_5\) values were below zero during most of the period considered, which means relation (3.5) was not observed again. Fig. 3.10 shows the portion of the trajectory that lies above the straight line (relation (3.5)). It also shows that during the summer period the trajectory oscillates about that same straight line. Additionally, in Fig. 3.9, it is also possible to observe an oscillatory behavior of \(L_5(t)\) after the new state of emergency was declared on November 9th, 2020, which may be related to increasing number of hospital beds made available by the NHS to respond to the new aggravation of the pandemic crisis that started early November and lasted until the end of January.
Figure 3.9: Selection of $\omega_4 = 0.5$ to get a trajectory of the linear combination $L_5(t)$ with values close to zero in the summer period.

Figure 3.10: Trajectory of the pair $(p_4, \frac{p_3 + p_5}{2})$ during the period under study.

In Fig. 3.10, the trajectory of $(p_4, \frac{p_3 + p_5}{2})$ is directed towards the origin $(0, 0)$ after January 31th, 2021, with the mean of the proportions of the worst pandemic events, $\frac{p_3(t) + p_5(t)}{2}$, rising above the level 0.5 by the end of February, 2021. In the middle of April, 2021 this mean and the infirmary occupancy rate were both lower than 20%, with the former higher than the latter.

As discussed previously, Figs. 3.8 and 3.10 each depicts similar shapes but with different ranges of values in the y-axis due to the values of death rate by COVID-19 ($p_3(t)$, illustrated in Fig. 3.8) being lower in magnitude than the utilization rate of ICU.

3.5 Conclusion

COVID-19 has been increasingly studied and scrutinized and although its long-term effects on individual health are still mostly unknown, its impact on public healthcare systems is very
strong and wide. In this work, five linear combinations and their corresponding MOVER-based interval estimates were built for extracting epidemiological information about the daily evolution of several interesting events related to the COVID-19 epidemic curve in Portugal, from which a new graphical technique for analyzing epidemiological data and, consequently, assessing the impact of either less or more restrictive measures that can be applied as prevention and control strategies over the spread of the disease was devised.

The joint analysis of the two kinds of graphs considered in this work showed an association between the implementation of public health measures and the easing of the pressure over the NHS. In the different graphs, it is possible to observe the “favorable situation” of the pandemic during the summer period, with the improvements being noticeable after the implementation of the first state of emergency that lasted 45 days, between March 18th and May 2nd, 2020. We can therefore infer that these measures helped stabilize the pandemic during the summer period. In the middle of September, 2020, a deterioration of the pandemic situation starts to be visible in the graphs, that is, Portugal started to move towards an unfavorable situation showing a trend for increasing the mean proportion of the worst conditions (hospitalizations and deaths by COVID-19) and having a high rate of nonhospitalized active cases. In October, 2020, the situation reached a level of increasing concern (due to the lack of resources of the NHS), reinforcing the need for the government and the NHS to increase the number of infirmary and ICU beds available for COVID-19 patients, as well as the number of specialized medical personnel. Those measures were applied on October 14th, but there was no visible improvement in the situation, which led to the implementation of a new, more restricted, state of emergency, that started on November 9th, 2020. Maximum levels of occupancy rates in both infirmary and ICU remained until the end of January, when the first begun to show signal of decrease (but not for the ICU occupancy rate) two weeks after the lockdown imposed on January 15th, 2021. As a consequence of this lockdown, the ICU occupancy rate begun also to decrease in February, achieving both a value lower than 20% in April, even after the removal some restrictions on March 15th, 2021.

The analysis of linear combinations of proportions of pandemic related events may be specially useful to monitor the evolution of the pandemic. By choosing appropriate proportions, these linear combinations can be used to characterize the level of utilization of the NHS resources. In the future, this new approach may be used in other epidemic outbreaks of communicable diseases. On the other hand, the analysis of these proportions is important for finding an appropriate model of low dimension, with a small number of compartments and parameters that can be estimated with good precision, which allows us to better understand the dynamic of the pandemics in terms of its impact over the resources of the NHS.
Acknowledgments

This work is supported by Portuguese funds through the CIDMA-Center for Research and Development in Mathematics and Applications, and the Portuguese Foundation for Science and Technology (FCT-Fundação para a Ciência e a Tecnologia), within projects UIDB/04106/2020 and UIDP/04106/2020. Cristiana J. Silva is supported by FCT via the FCT Researcher Program CEEC Individual 2018 with reference CEECIND/00564/2018. Juliana Reis was supported by a FCT research grant with reference BI-BIOMATH-2-04/2020 in the scope of the Thematic Line BioMathematics of CIDMA.

References

[1] Diário da República, Decreto do Governo que regulamenta o estado de emergência, Decreto n. 2-a/2020, Available at https://dre.pt/home/-/dre/130473161/details/maximized, 2020. (Accessed 1 November 2020).
[2] Diário da República. Regulamenta o estado de emergência e o estado de calamidade para o período entre 1 e 3 de maio de 2020, Decreto n. 2-d/2020, Available at https://dre.pt/home/-/dre/132883343/details/maximized, 2020. (Accessed 1 November 2020).
[3] Diário da República, Renovação do Estado de Emergência, Available at https://covid19estamoson.gov.pt/renovacao-do-estado-de-emergencia/, 2020. (Accessed 18 November 2020).
[4] Direção Geral de Saúde, Capacidade de medicina intensiva aumentou, Available at https://covid19.min-saude.pt/capacidade-de-medicina-intensiva-aumentou-23/, 2020. (Accessed 30 November 2020).
[5] Direção Geral de Saúde, Ponto de Situação Atual em Portugal, 2021, Available at https://covid19.min-saude.pt, 2020. (Accessed 17 September 2021).
[6] RTP Noticias, Available at https://www.rtp.pt/noticias/pais/foram-ultrapassadas-as-camas-destinadas-covid-19-em-uci_v1273788, 2020. (Accessed 30 November 2020).
[7] COVID-19 situation update worldwide, week 36 of 2021, European Centre for Disease Prevention and Control, Available at https://www.ecdc.europa.eu/en/geographical-distribution-2019-ncov-cases, 2021. (Accessed 17 September 2021).
[8] Data Science for Social Good Portugal, Available at https://github.com/dssg-pt/covid19pt-data/blob/master/data.csv, 2021. (Accessed 16 April 2021).
[9] eVM - Vigilância de Mortalidade, Available at https://evm.min-saude.pt/#shiny-tab-q_total, 2021. (Accessed 16 April 2021).
[10] A. Agresti, B.A. Coull, Approximate is better than “exact” for interval estimation of binomial proportions, American Statistician 52 (2) (1998) 119–126.
[11] S. Escudeiro, A. Freitas, V. Afreixo, Approximate confidence intervals for a linear combination of binomial proportions: a new variant, Communications in Statistics. Simulation and Computation 46 (9) (2017) 7501–7524.
[12] G. Giordano, F. Blanchini, R. Bruno, P. Colaneri, A. Di-Filippo, A. Di-Matteo, M. Colaneri, Modelling the COVID-19 epidemic and implementation of population-wide interventions in Italy, Nature Medicine 26 (2020) 855–860.
[13] D.S. Hui, E. Azhar, T.A. Madani, F. Ntoumi, R. Dar, G. Ippolito, T.D. Mchugh, Z.A. Memish, C. Drosten, A. Zumla, E. Petersen, The continuing 2019-nCoV epidemic threat of novel coronaviruses to global health—the latest 2019 novel coronavirus outbreak in Wuhan, China, International Journal of Infectious Diseases 91 (2020) 264–266.
[14] S.M. Kissler, C. Tedijanto, E. Goldstein, H.Y. Grad, M. Lipsitch, Projecting the transmission dynamics of SARS-CoV-2 through the postpandemic period, Science 368 (6493) (2020) 860–868.
[15] H. Li, S.M. Liu, X.H. Yu, S.L. Tang, C.K. Tang, Coronavirus disease 2019 (COVID-19): current status and future perspective, International Journal of Antimicrobial Agents 55 (5) (2020) 105951.
[16] A. Martin-Andrés, M. Álvarez-Hernández, I. Herranz-Tejedor, Inferences about a linear combination of proportions, Statistical Methods in Medical Research 20 (4) (2011) 369–387.
[17] A. Martín-Andrés, I. Herranz-Tejedor, M. Álvarez-Hernández, The optimal method to make inferences about a linear combination of proportions, Journal of Statistical Computation and Simulation 82 (1) (2012) 123–135.

[18] F. Ndairou, I. Area, J.J. Nieto, D.F.M. Torres, Mathematical modeling of COVID-19 transmission dynamics with a case study of Wuhan, Chaos, Solitons and Fractals 135 (2020) 109846.

[19] R.G. Newcombe, Interval estimation for the difference between independent proportions: comparison of eleven methods, Statistics in Medicine 17 (8) (1998) 873–890.

[20] S. Wallis, Binomial confidence intervals and contingency tests: mathematical fundamentals and the evaluation of alternative methods, Journal of Quantitative Linguistics 20 (3) (2013) 178–208.

[21] E.B. Wilson, Probable inference, the law of succession, and statistical inference, Journal of the American Statistical Association 22 (158) (1927) 209–212.

[22] G.Y. Zou, W. Huang, X. Zhang, A note on confidence interval estimation for a linear function of binomial proportions, Computational Statistics & Data Analysis 53 (4) (2009) 1080–1085.