Regular Article

THE EUROPEAN PHYSICAL JOURNAL PLUS



A note on Corona-virus (Covid-19) Infection in Italy

Angelo Pagano^{1,a}, Emanuele V. Pagano^{2,b}, Grazia Pagano^{3,c}, Santi Spartà^{4,d}

¹ Istituto Nazionale di Fisica Nucleare (INFN), 64, via S.Sofia, 95123 Catania, Italy

² National Laboratories of Sud (INFN-LNS), 62, via S.Sofia, 95123 Catania, Italy

³ Dipartimento di Matematica e Informatica Università di Catania, 64, via S.Sofia, 95123 Catania, Italy

⁴ Radiation and Robotic s.r.l., 25, Via Chianello, 95024 Acireale, Italy

Received: 7 July 2020 / Accepted: 6 August 2020 / Published online: 18 August 2020 © Società Italiana di Fisica and Springer-Verlag GmbH Germany, part of Springer Nature 2020

Abstract A statistical analysis of the corona-virus (Covid-19) infective process has been performed by a cooperative action during the period February-June 2020. A good analysis has been obtained by using an entropic model typical of phenomena where statistical entropy-negaentropy balance is expected to play a major role. A saturation value of the infected humans was observed, and the number of people potentially (asymptomatic) involved in the process was determined with an accuracy of 15% in the Italian case, as relevant example. The saturation value represents about 16% of the total (symptomatic + asymptomatic) involved population in the process. The stability of the observed saturation level with the time shows that the governmental lockdown prescriptions, guided by scientists (virologists) have been effective to contain the diffusion of the virus and the associated human mortality.

1 Introduction

The total number of Covid-19 infected humans in Italy in the period February-June 2020 has shown an increasing behaviour, with the time typical to the one expected in infectious Diseases of Humans [1]. Alert has been promptly evidenced by scientists after the first events in China [2,3]. Monitoring this number by a suitable statistical analysis [4–7] is an essential step in order to understand the phenomenon with the purpose to disentangle among possible existing models [8] and, consequently, to achieve a phenomenological description of the dynamical process [9–11]. In the simplest virological model [8], the *S1S* model, it is assumed that each individual could be infected and recovered, following the transitions: $S \rightarrow I \rightarrow S \rightarrow ...$, for an illimited number of times, where *S* is the class of "non-infected" individual and *I* is the class of infected individuals. In this model, it is highly desirable to obtain quantitative indications about two phenomenological main parameters: β and γ . The parameter β is the transmission coefficient of the infective process, following the symbolic transition: $S \rightarrow I$; it depends on the socio-economic structure and the intrinsic nature of the virus. The inverse of the second parameter: $\frac{1}{\gamma}$ gives the average time of permanence of

^a e-mail: angelo.pagano@ct.infn.it (corresponding author)

^be-mail: epagano@lns.infn.it

^c e-mail: grazia.pagano89@gmail.com

^de-mail: radiation.robotic@gmail.com

an individual in the class I, following the symbolic transition: $I \rightarrow S$. They enter in the time-dependent differential equation Eq. (1), describing the time evolution of the a priori probability p (i.e. given by the ratio between infected individuals and the total population) associated to the infectious process [8]:

$$\dot{p} = (\beta - \gamma)p - \beta p^2 \tag{1}$$

The stationary solution of Eq. (1) is asymptotically stable $(p \sim 1 - \frac{\gamma}{\beta})$ if $\beta \geq \gamma$ (plateau). If $\beta \leq \gamma$, the epidemic process goes through a natural extinction. Extensions of *SIS* model should require a diffusion term: an infective process originated in a well-defined position is supposed to expand tendentially in space (similar to a propagation wave-phenomenon) involving larger areas and a larger number of potentially infected individuals [11,12]. The solving process of Eq. (1) has to be performed by taking into account for large medical expertises and detailed knowledges of the time evolution of the infective process. However, as general rule, the simple model in Eq. (1) is too crude in real case, as the one observed recently in Italy and much more complicated model should be used, in particular by systems of coupled differential equations (neural network) [8, 13].

For practical purposes, it is more convenient to adopt a less powerful method, but useful to predict, in real time, few important statistical global parameters, such as the number of expected infected humans, the growing up of the infection with the time, the velocity of the propagation of the infection, etc..., in order to adopt prompt and reasonable anti-virus prescriptions aimed at reducing the effect of the infection among the population.

In our approach, we follow a method (differential equation Eq. (2)) that is common in any entropic process: from the behaviour of complex physical systems [14] to the description of macro-economics processes [15]. The model, under reasonable assumptions, could be applied just at the beginning of the process (initial conditions t = 0) to predict future probabilistic time evolution. It has to be noted that the accuracy of the predictions is dependent upon the adopted data sampling method at the beginning of the infective process (corona-virus swab methods).

The entropic model used in the analysis shown below, when applied to Italian case, predicts a saturation level of the infected humans around 240.000 units. Our results suggest that the entropic-model (curves red-green-black in Figs. 1, 2 and 3) represents significantly the statistical dynamical evolution of the infection process.

2 Data analysis and modelling

A statistical analysis of Covid-19 data (from February 20th to the end of June 2020) by adopting an Entropic model has been performed following the time evolution of the infective process [16]. In that model, the cumulative number Y(t) of infected humans as a function of the time is one of the solutions of the differential equation:

$$\dot{Y} = b \times Y - c \times Y^2. \tag{2}$$

Remarkably, the formal expression of Eq. (2) is the same as in Eq. (1).

The unique solution of the differential equation Eq. (2) is determined by evaluating the three parameters: *K*, *b* and *c* in the functional form:

$$Y(t) = \frac{Kb \times e^{bt}}{Kc \times e^{bt} + 1}$$
(3)

The three parameters have been extracted by fitting the experimental data of the daily rate of infected humans (see Fig. 2) at early stage of the emerging infection by using the first derivative of formula Eq. (2), i.e. the expression:

$$\dot{Y}(t) = \frac{Kb^2 \times e^{bt}}{(Kc \times e^{bt} + 1)^2}.$$
(4)

In our analysis, this fitting procedure was done in a period of about 30 days (from February 20th 2020 to March 20th 2020).

The parameter *b* measures the probability of the infection process for unitary time (day). And, evidently, 1/b measures the average time, τ , of the permanence of an individual in the class *I*. The parameter *c* roughly quantifies the empirical entropic-negaentropic balance (average equilibrium between the number of new infected and the number of healed).

For the fitting procedure, we adopted the ROOT software [17], well known to experimental nuclear physicists, incorporating the MINUT routine for Least Squares minimisation with respect to the parameters and error-matrix accuracy computation. Evidently, due to quasi-exponential behaviour of the data (see Fig. 1) at early stage of the emerging infection, the parameters K, b and c have different sensitivities, and consequently, they have shown different errors, ranging from 10% (b) to 40% (c). However, the crucial parameter b - used for N_0 and - τ determinations, shows a reasonable accuracy, essentially determined by the statistics of the data sampling method.

From the parameters *K*, *b* and *c*, the total number, N_0 , of the population statistically involved in the infection (reservoir) at the beginning of the emerging infection (t = 0) and the value τ have been evaluated. So, the solution of Eq. (2) is given by formula Eq. (3) with the following parameters :

$$K = 1916(\pm 661)day, \quad b = 0.17(\pm 0.012)day^{-1}, \quad c = 8.010^{-7}(\pm 3.010^{-7})day^{-1}$$

(See the insert in Fig. 1). The daily rate of infected humans, as given by formula Eq. (4), is shown in Fig. 2. (See the insert in Fig. 2). The curves reported in Fig. 2 show the typical behaviour of the "logistic function", well known to the economists, with the characteristic "rise" (negaentropic phase) and "fall" (entropy phase) behaviours. Usually, the entropy phase indicates a strong economic crises with drastic reduction of the production of consumer goods and financing activities [14]. In our case, the displacement of the infection from the north to the south of Italy (see below for discussion), modifies the simple quasi-gaussian tail of the entropy phase producing a nearly exponential tail smoothly decreasing with the time.

The daily relative rate as evaluated by the ratio of the daily infected (Eq. (4)) over the number of cumulative infected humans (Eq. (3)) is given by:

$$\frac{\dot{Y}(t)}{Y(t)} = \frac{b}{Kc \times e^{bt} + 1} \tag{5}$$

(See the insert in Fig. 3)

After the first determination of the parameters K, b and c no further adjustments have been made. As a consequence, a satisfactory agreement of the model (Figs. 1, 2 and 3) with the experimental data as a function of the time is deduced. In Figs. 1, 2 and 3 three different curves (with different colours) are there shown. They represent, respectively, a lower limit evaluation-black one-, an upper limit evaluation-green one- and the most probable evaluationred one-(see figure caption of Fig. 1 for details).

3 Discussion and conclusion

The number of experimentally determined infected (data in blue symbol) is significantly lower than the most probable predictions (red line) in the range [35-60] days evaluation, indicating a steady displacement of the infective process with the time along with the territory. This effect indicates a regional displacement of the infection from the north to south of Italy. As a consequence, starting from 34-day evaluation, a nearly linear dependence of the total number of infected as a function of the time has been observed. This behaviour could be interpreted as a superposition of different logistic functions steady moving from the left to the right with the time (Fig. 1).

The displacement velocity of the infection ranges between 25 and 50 km/day during first 35-70 days evaluations, achieving a stationary behaviour (and a vanishing displacement), due to the effects of the lockdown restriction. The agreement between the predictions at the saturation value and the experimental data is quite good within the statistical accuracy (given by the parameter-variances σ of the fitting procedure). As new suggestion, we note that the value of infected humans at the "saturation level" (around 240.000) has been determined already at the beginning of the infection (t=0) and that the time evolution of the process (see Figs. 1, 2 and 3) is given mainly by the particular adopted daily corona-virus swab methods in Italy. The measured (apparent) increasing behaviour of the infection with the time (Fig. 1) has been determined by three main factors:

- (i) the adopted data sampling method (corona-virus swab methods) due to a daily sampling (from a fixed reservoir) at nearly constant daily rate (i.e. by changing the method a slightly different behaviour should be observed).
- (ii) the displacement from the north to the south of Italy in the early phase (before the lockdown) of the infection.



Fig. 1 Integrated Infected humans as detected by corona-virus swab (Italian) methods as a function of the time. Different curves indicate statistical significance: most probable (red), adding the errors in the parameters (black), subtracting the errors in the parameters (green). The read curves indicate saturations (asymptotic) behaviours. The number of experimental determined infected humans (blue symbols) is significantly lower than the most probable predictions (read) in the range: [34-60] days evaluation, indicating a steady displacement of the number of detected infected with the time



Fig. 2 Rate of infection as a function of the time (derivative of function in Fig. 1) compared with experimental data. Different curves have the same meaning as described in Fig. 1



Fig. 3 Relative rate of the infection as a function the time. The function (red one) reproduces the history of the infection in very good agreement with the data within the experimental errors (black and green functions). Different curves have the same meaning as described in the caption of Fig. 1

(iii) Lockdown restrictions since March 10th 2020.

So, we conclude that the saturation value has an intrinsic virological value related to the infection; while the raising measured behaviour (Fig. 1) is affected by different accidental factors (i-ii-iii).

As main numerical results of our fitting procedure, it is argued that the infectious process has involved (within the considered period of time) a number of about $N_0 \sim 1.5$ millions

humans in Italy (symptomatic + asymptomatic), with an accuracy of about ~ 15% and an average time of permanence, τ of about 6 days (order of magnitude). The N_0 value is estimated (in agreement with our assumption) by assuming that the saturation value (symptomatic ~ 240.000) was given by the approximation (instantaneous process) : $\delta N = N_0 \lambda \delta t$ ($\lambda = b$ and $\delta t = 1$ day). So, our analysis indicates that the number of infected at the saturation represented a fraction ~ 4/1000 of the total Italian population and ~ 16/100 of the involved population in the infective process (symptomatic +asymptomatic = 1.5 million). It has to be noted that most of the involved individual (practically, $\geq 60\%$) were localised in the north-regional countries.

The stability of the saturation level shows that the governmental lockdown prescriptions, guided by scientists (virologists) have been effective to contain the diffusion of the virus and the associated human mortality. The human mortality in Italy being among the highest value in the world has not explanation in the frame of our simple statistical model. A complete explanation requires extensive medical and genetics studies and, consequently, is out of the present approach.

Acknowledgements The authors are grateful to Elena Geraci for reading a preliminary version of the draft.

References

- 1. D.M. Morens, G.K. Folker, A.S. Fauci, The challenge of emerging and re-emerging infectious diseases. Nature **439**, 242 (2004)
- C.I. Paules, H.D. Marston, A.S. Fauci, Coronavirus infections-more than just the common cold. JAMA 323(8), 707 (2020)
- A.S. Fauci, H. Clifford Lane, R. Redfield Robert, Covid-19: navigating the uncharted. New Engl. J. Med. 372, 1268 (2020)
- I. Ciufolini, A. Paolozzi, An improved mathematical prediction of the time evolution of the Covid-19 pandemic in Italy, with a Monte Carlo simulation and error analyses. Eur. Phys. J. Plus 135, 495 (2020)
- M. Radiom, J.-F. Berret, Common trends in the epidemic of Covid-19 disease. Eur. Phys. J. Plus 135, 517 (2020)
- 6. A. Bonasera, S. Zhang, Chaos, percolation and the coronavirus spread. Front. Phys. 8, 171 (2020)
- P. Castorina, A. Iorio, Data analysis on Coronavirus spreading by macroscopic growth laws. Int. J. Mod. Phys. C 31(7), 2050103 (2020)
- C. Piccardi, Reti Sociali e Diffusione di Epidemie, Lettera Matematica, 86 Pristeam Univ (Springer, Bocconi, 2013), p. 30
- R. Anderson, R. May, Infectious Diseases of Humans: Dynamics and Control (Oxford University Press, Oxford, 1991)
- 10. R. Anderson, *The Application of Mathematical Models in Infectious Disease Research*. http://www.ph.ucla.edu/epi/faculty/olsen/200B2010/ANDERSON.FITL.2001.pdf
- A. Banos, J. Lacasa, Spazio-temporal exploration of SARS epidemic. Cybergeo: Eur. J. Geogr. [on line], http://cybergeo.revue.org/12803) (2007)
- R. Pastor-Satorras, A. Vespignani, Epidemic spreading in scale-free networks. Phys. Rev. Lett. 86, 2909 (2001)
- 13. M.E.J. Newman, The structure and function of complex networks. SIAM Rev. 45, 167 (2003)
- 14. N. Georgescu-Roegen, *The Entropy Law and the Economic Process* (Harvard University Press, New York, 1999)
- G. Amata , S. Notarrigo , Energia e Ambiente: una ridefinizione della teoria economica. C.U.E.C.M. Catania, Italy (1987)
- COVIDSTAT-INFN, Monitoring Covid-19 emergence data. https://www.openaccessrepository.it) DOI Badge 10.15161/oar.it/30431 (2020)
- R. Bruna, F. Radernaker, ROOT: an object oriented data analysis framework. Nucl. Inst. Methods Phys. Res. A 389, 81 (1997)