

## Editorial

---

# Statistical modeling of pandemics and coronavirus

Igor Mandel<sup>a</sup> and Stan Lipovetsky<sup>b,\*</sup>

<sup>a</sup>*Redviser Inc., NJ, USA*

<sup>b</sup>*MASA Co-Editor-in-Chief, MN, USA*

This special MASA issue is intended for the problems of statistical modeling of pandemics in general, and the Coronavirus COVID-19 one particularly. A recent analysis in *Nature*<sup>1</sup> shows that the number of papers on coronavirus skyrocketed in the first 4 months of 2020 and then stabilized, more or less in accordance with behavior of the pandemic itself. The statistical modeling of the coronavirus pandemic is also flattened – yet the number of monthly publications is huge, exceeding the “normal pre-pandemic” level in 25–30 times.

The goal of “modeling” is to create multiple scenarios, including the pessimistic and optimistic, to be immediately available when circumstances require it. Perhaps, the reasons for the pandemic been so devastating are that the science was not ready, the WHO recommendations were not in the place, effective government plans did not exist, and so on. The period of the preliminary preparations was just lost, which is especially sorrowful, because comparatively recent pandemics, like SARS in 2002–4 and others, gave all the reasons to be timely prepared. It seems, just Taiwan<sup>2</sup> took all previous cases seriously and made a strategic plan, which was brazenly ignored by other countries and WHO; the difference between Taiwan and other countries outcomes is now startling.

In light of that all, what could be the purpose for the special issue of the statistical journal on pandemic problems? It obviously will not help to reach the ear of decision makers in the struggle with the current wave, which seems starts to calm down. However, the different approaches presented in this issue will help in future preparation for the yet unknown pandemics or epidemics. A wide geography of the authors’ countries and variety of the topics cover somewhat different aspects of statistical modeling of pandemics. A reader should also know that all the papers were in preparation for several months earlier to this issue, while the pandemics was evolving very fast. Some of the quantitative results may look obsolete (although, the authors tried to get maximum in their data collection), but the methodological value of the proposed approaches stays to be useful.

Fighting the Coronavirus COVID-19 pandemic required quick developing tests and vaccines, continuing trials and research (Mandel & Lipovetsky, 2020). As a reflection of these efforts, multiple journal articles have been published on the related topics. The coronavirus pandemic covers the most populated areas on Earth, and the spread of infection has been going fast with the global transportation and connectivity of travelers and commerce. With COVID-19 highly infectious features, high transmissivity, often asymptomatic appearance, it spreads with huge consequences in areas of dense populations and poor public health systems. In conditions of the lack of a vaccine, only the forced isolation of the infected serves to decreasing the infection rates. However, within months of the virus

---

\*Corresponding author: Stan Lipovetsky, MASA Co-Editor-in-Chief, MN, USA. E-mail: stan.lipovetsky@gmail.com.

<sup>1</sup>How a torrent of COVID science changed research publishing – in seven charts, *Nature*, Dec 16, 2020, <https://www.nature.com/articles/d41586-020-03564-y>.

<sup>2</sup>COVID-19: Virus ‘success’ Taiwan to keep restrictions despite vaccine, Dec 11, 2020 <https://www.bbc.com/news/world-asia-55269729>.

appearance in China, the modern science succeeded in decoding its' genome, tracking its motion through the world, finding its genetic variations and evolution, and developing therapeutic treatments for the infected and the effective vaccines. This way to eliminate a human pathogen would lead the world to recovering from the pandemic itself and the economic disruptions caused by the pandemic. Effects of this pandemic and its impact on humanity at large would define the unknown features of living in the twenty-first century, with needs of global cooperation on solving the problems which have no national borders. And here we see our contribution to the whole mankind efforts to survive the pandemics with the tools which address this crisis and can help to cure in the post-pandemic world as well.

The special issue consists of the following works.

The paper by V. Moreau (Brazil) *Using the Weibull Distribution to Model COVID-19 Epidemic Data* shows that this distribution could quite successfully describe the appearances of cases and deaths in different countries. It is especially important because it allows to work with 2-waves situation which is observed now, if the transition moment was captured. The quality of approximation varies from country to country, but in general for majority of cases it is acceptable.

The work *A Probabilistic Epidemiological Model* by H. Switkay (USA) analyzed practical applications of the Beta-family distributions. The author challenges the traditional for epidemiology compartmental approach and offers a probabilistic model which outperforms the common method for several countries.

*A Hybrid of Artificial Neural Network, Exponential Smoothing, and ARIMA Models for COVID-19 Time Series Forecasting* paper by S. Safi and O. Sanusi (United Arab Emirates) is a non-typical example of merging of several traditional methods for time series with neural network in one model. Authors demonstrate that this combination provides better results than a single of these methods and made successful forecast of main pandemics indicators.

*Social Network Analysis and Community Detection on Spread of COVID-19* paper by A. Wickramasinghe and S. Muthukumarana (Canada) suggests mapping of distribution of the coronavirus bearers among the countries and by Canadian provinces to identify the main centers of pandemic, with a good validation of the model. This approach could be used on the early stage of a pandemic to set up priorities for banning the countries or provinces for travelers, if necessary.

The paper *Modeling COVID-19 Positivity Rates and Hospitalizations in Texas* by R. Kafle et al. (USA) addresses an important question of a change in direction of a trend or slope in the data. The authors use the effective Bayesian joint-point regression technique, analyze the indicator of the level of hospitalization and demonstrate that it could be successfully predicted. What is especially important – they even could capture the beginning of the second wave of the virus.

In a similar vein, the paper *Statistical Methods for Estimating Cure Fraction of COVID-19 Patients in India*, by E Sreedevi and P. Sankaran (India), considers a fraction of the cured people in the country and a way to model it. They used different survivor functions, including Kaplan-Meier mapping, and propose a specific way for estimation of the model parameters. The predicted probability of surviving for different demographic groups matches the data quite well.

A specific approach is presented in the paper *Spatial Identification of Component-based Relative Risks* by S. Kadhem (Iraq), which considers the Poisson mixture distribution applied to different regions and shows that risk estimated by those models fits the observed rates much better than more traditional estimates based on standardized mortality ratio. The model has been successfully tested in 18 Iraqi regions.

Another paper by V. Moreau (Brazil) *Inconsistencies in Countries COVID-19 Data Revealed by Benford's Law* covers a delicate and controversial topic of reliability of official pandemic data in 80 countries. The author has checked the distribution of the first digits of the new cases and concluded that for many countries that distribution does not fit well the Benford's law, at least based on the criteria used (p-value less than 0.01). Of course, such findings do not directly mean the fraud, but they hint on the quality of data or a need in additional deeper testing. The USA enters the same group with Russia, Iran, and Belarus, which raises a question about the data quality, see also (Mandel & Lipovetsky, 2020; Forsberg, 2021).

Besides these papers comprising the special issue on pandemics, there is an additional note in this journal issue with the comments on the paper (Mandel, 2020) devoted to some newly appeared ethical problems in science, particularly, in our area of statistics, and concerning such great figures in statistics as its founders R. Fisher, K. Pearson, and F. Galton. All the readers are invited to share their opinions on this problem.

## **References**

- Mandel, I., & Lipovetsky, S. (April 8 2020). COVID-19: Can we estimate infection speed and fatalities? *American Thinker*, [https://www.americanthinker.com/articles/2020/04/covid19\\_can\\_we\\_estimate\\_infection\\_speed\\_and\\_fatalities.html](https://www.americanthinker.com/articles/2020/04/covid19_can_we_estimate_infection_speed_and_fatalities.html).
- Forsberg, O.J. (2021). *Understanding Elections through Statistics: Polling, Prediction, and Testing*. CRC, Boca Raton, FL.
- Mandel, I. (2020). The science of statistics in victimhood culture: Dignity and dishonor. *MASA*, *15*(4), 181-195.