Statistics for social good: trying to model better coronavirus predictions

Logistic growth models might help with COVID-19 forecasting, say professors Jeroen Neckebrouck and José Azar. Here's a look at a side project they launched to help get a handle on the public health crisis.

March 27, 2020

- **Professors Jeroen Neckebrouck and José Azar are trying to forecast the spread of the novel coronavirus based on insights from logistic growth**

models.

- **Population growth models are used because they precisely predicted the evolution of confirmed infections in China and South Korea.**
- **According to the model, the daily increase in the number of confirmed cases in Italy and Spain should continue to diminish in the week ahead.**

Like so many, Jeroen Neckebrouck, professor of entrepreneurship at IESE, has followed news of the coronavirus pandemic with alarm. But with his knowledge of statistics, he was also occasionally upset by some predictions he was reading in the media.

He teamed up with his colleague José Azar, professor of economics, with the idea of helping with forecasting as a public service. "We're obviously not medical doctors, we can't help out on the frontlines, but we do know statistics and econometrics, so we can try to put our skills to use otherwise," Neckebrouck says.

To do this, they turned to population growth models, as put forward by Edwin van den Heuvel, Marta Regis and Zhuozhao Zhan from Eindhoven University of Technology. They found that the COVID-19 outbreaks in China and South Korea followed the S-shaped curve seen in population growth -- with a slow start, rapid uptick and then a curve that levels out once it nears a maximum -- in a way that could be predicted by the model with 98-99% accuracy.

Would hotspots in Europe, North America and elsewhere follow the same pattern? Based on the shape of early data, could contagion predictions be improved? For Neckebrouck and Azar, it was certainly worth looking into. To advance the previous work, they calibrated their models for estimating bias and cleaned up the data, which seems to have improved predictions and enabled them to make forecasts for up to seven days into the future, within confidence intervals of 68% and 90%.

The pace of the outbreaks in Italy and Spain, for example, have surprised many, and yet, as more daily data is collected, it looks increasingly likely that the rate of growth is slowing down for the remainder of March. The authors emphasize that they are not epidemiologists, and that the statistical models do not take into account government measures nor cultural differences in behavior from country to country. "But so far, we've seen most of the actual number of cases reported fall within our forecast range," Azar summarizes.

Each morning, Neckebrouck collects data from the Johns Hopkins University-hosted dashboard, checks it against other reputable sources, and tries to make improvements to his

site as Spain remains in lockdown.

Stay tuned to see how the current predictions fare over time. "Comments and suggestions for improvements are welcome," he remarks. "And please remember to stay inside." Visit: https://sites.google.com/view/corona-predictions