Understanding the COVID-19 Model and How it was Used

Thomas R. Cuba, 2020 0410

There are several types of models. Let's take ones that we are all familiar with in order to make the point. Every American has seen at least two models: the linear model we all learned about in 5th grade and the spaghetti models used to predict the tracks of hurricanes. Understanding these two, we can understand the CDC's COVID-19 model and how it was used.

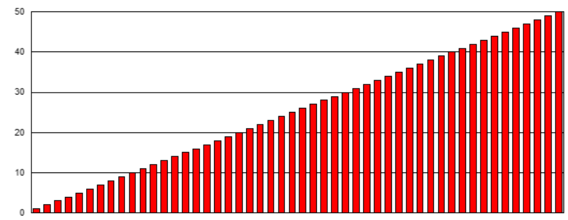

Figure 1: y = (A*x)+B

In the 5th grade we all learned that two variables (x and y) can have a relationship that is a straight line when plotted on graph paper. The equation (or formula) for the relationship is y=(A*x)+B, where the values of A and B are not, in the generic form of the equation, known, but are always the same for every value of x. Because the values of A and B are always the same in the relationship, they are called "Constants." The relationship results in the graph shown in figure 1.

When used in an application, such as estimating how much gas might be needed to drive from one place to another, a value for A and B are chosen by observing the gas consumption for the car being driven. Then, for any number of miles driven (represented by x), the predicted number of gallons of gas (y) can be calculated based on the relationship.

Even in complex situations such as aiming the Apollo spacecraft well enough to reach the moon several days after launch, the model of where the moon would be is very exact because the "Constants" are very well known, even though there are many, many of them and not just A and B.

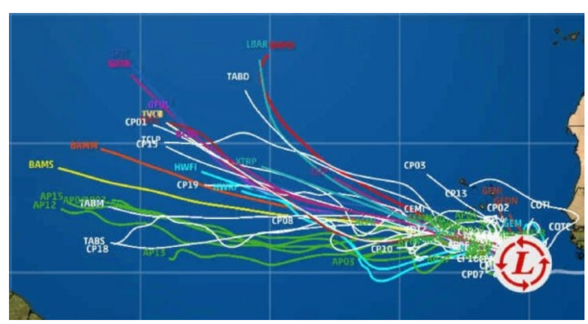

Figure 2: Spaghetti models.

In other situations, the constants aren't as well known and the values aren't always the same throughout the model. The most recognizable examples of this type of model are the Hurricane Prediction Models. There are so many different methods of preparing these that the weather service routinely presents what is called the Spaghetti Models as shown in figure 2. In the figure, each line represents a different model in which different mathematicians have used different relationships for the data. But, the most important thing to know is that most of them use the <u>same</u> data. The mathematicians just put the data together differently in their version of the model.

When the weather service finally picks a single model to use for the actual forecast, the graphic that is used will include only one line, but that line will be surrounded by what is known as the "Cone of Uncertainty." What the cone represents brings us back to the values of A and B.

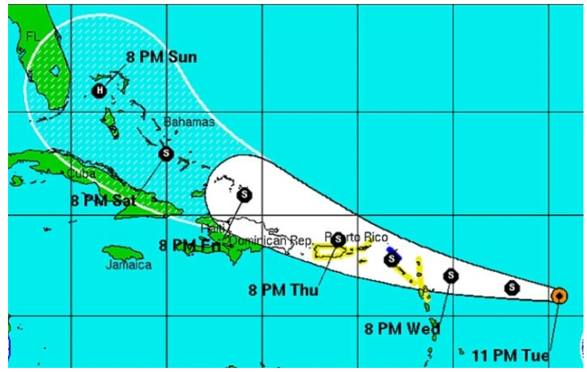

Figure 3: The Cone of Uncertainty.

In most situations, even though they are called Constants, the values of A and B are values that lie between a high and low observed value. In figure 4, the values of both A and B have been selected randomly (but within limits equal to the high and low of the range) to reflect the possible range of values. The general shape of the line strongly resembles that in figure 1, but the differences between the higher values and lower values of y are clear. These are the edges of the Cone of Uncertainty for this example. In figure 4, the values of A and B used in each calculation of y have been given more possibilities and, while the general trend is still similar to figure 1, the uncertainty increases and the reliability decreases.

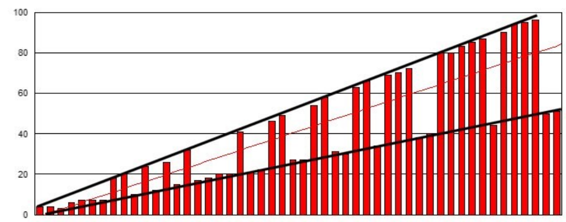

Figure 4: values of A and B are not discrete but occur within a range. The black lines are equivalent to the uncertainty.

In more complex models, sometimes a model is used to generate each value of A and B that might be used to calculate each value of y. In figure 5, A and B have been given the ability to oscillate between high and low values on a random basis. The Cone of Uncertainty increases wildly.

That's enough about the mathematics of models. Now let's consider how they are used. When a model has a Cone of Uncertainty, as in the Hurricane, many people will also assign a probability to the centerline of the graphic. This is generally wrong, because the probabilities are what have defined the edge of the cone in the first place. The hurricane can make landfall anywhere within the cone.

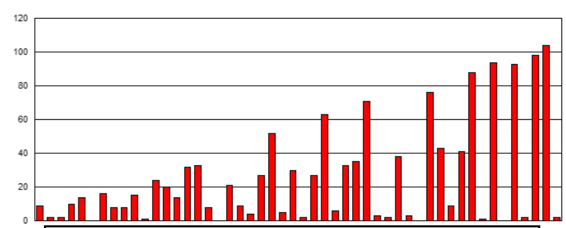

Figure 5: values of A and B are computed mathematically including an oscillation.

In many situations where these types of models are used, the general public, or consumer, has a desired outcome. Farmers in Iowa want or desire a certain amount of rain. If the model predicts flooding, that is often referred to as a "Worst Case Situation."

The Worst Case Situation is constructed by the mathematician in order to demonstrate how bad things can get. In doing so, the mathematician will select values for the Constants in the model which are least desirable.

Turning to the COVID-19 model, if the mathematician chooses a value that indicates that 5 out of 100 people who catch the disease will die, the outcome is much worse than if he chooses a value that indicates that 1 out of 100 people who catch the disease will die. Choosing the value 5 represents a worst case. In the same manner, if the mathematician decides that 9 out of 10 people who come within 6 feet of an infected person will catch the disease, that is a worse case than if he decides that 1 out of 10 people will catch it under the same circumstances. These two examples also demonstrate the geometric relationship among choices. The force of the first choice is multiplied by the force of the second choice. If more people catch the disease then many more people will die.

Even in the Worst Case Situation, the model has a Cone of Uncertainty. When the COVID-19 predictions were first released there were many models from many sources and the predicted deaths in the United States ranged from around 21,000 (best case) to 2.2 million (worst case).

What the CDC presented in February was the Worst Case situation and they did not show the Cone of Uncertainty, so we don't know if the 2.2 million deaths originally predicted was the centerline of the model or the upper edge: the worst of the worst case.

What the CDC did was to present the Worst Case as the "Most Probable" case. The faith and trust that the administration and the people had in the CDC led us all to enter into a state of fear never before experienced in this nation, and in fear, we relinquished many of our freedoms.

As of the date of this paper, the predicted number of deaths has been reduced to 60,000 and social distancing has been hailed as the causative agent for the reduction. The claim is based on another model with multiple variables and multiple values for constants used in the calculations.