Estimating the death toll of Hurricane Maria

Puerto Rican authorities report only 64 deaths from Hurricane Maria, the worst storm to hit the island in nearly a century. But statistical estimates point to many more victims. By **Roberto Rivera** and **Wolfgang Rolke**

On 20 September 2017, Hurricane Maria made landfall in Puerto Rico with sustained winds of 249 kilometres per hour, bringing with it at least 400 mm of rain across much of the island and leaving residents without power, running water or any form of communication. Three months after the event, the power authority was generating only 68% of what it produced before the storm, and it was unclear how many of the island's 3.35 million residents were still struggling without electricity.

The severity of the storm posed a serious risk to life, both during the event and in its aftermath. It was particularly dangerous for the sick and elderly, who may have been cut off from vital aid (nyti.ms/2krJrAV). Based on forensic evaluations, the Puerto Rican authorities report that the storm caused 64 deaths. Unfortunately, forensic examination becomes unreliable when a storm like Maria devastates a country's infrastructure. In addition, it is problematic to determine if some deaths, say those caused by strokes and diabetes, were indirectly related to the storm – and indeed the bodies of hundreds of people who died of supposed "natural causes" were cremated without proper forensic examination (bzfd.it/2iY64MY).

Local media outlets (bit.ly/2hTUhvy) have suggested that the true death toll of Maria is much higher than 64. CNN estimated in November that Hurricane Maria caused 499 deaths in the first month after landfall (cnn.it/2izHkqY). In December, the *New York Times* estimated 1052 additional deaths from 19 September until 31 October 2017, as compared to previous years (nyti.ms/2nJgDoH).

Estimating excess deaths due to natural disasters is an important problem to address. Families are eligible for US federal assistance for deaths attributed to a major disaster or emergency, while the death toll tells other countries about the severity of the situation and whether aid is needed (cnn.it/2izHkqY). Most importantly, residents can make informed and potentially life-saving decisions when they know the risks they face.

The question we want to answer is: are the 64 deaths officially attributed to Hurricane Maria compatible with mortality data before the atmospheric event? According to mortality data from the Demographic Registry Office (DRO), 1582 deaths were recorded in Puerto Rico from 1 September to 19 September 2017 (the day before Maria made landfall), and there were 1317 deaths

Year	Deaths
2010	77.85
2011	79.38
2012	79.50
2013	79.79
2014	82.56
2015	76.22
2016	77.29
2017	102.83

TABLE 1 Average daily deaths in September and October: total deaths over 61 days. For 2017, the average is from 20 September to 31 October.

Our model

It can be shown using a binomial model that the maximum likelihood estimate of the Maria death toll is a function of the difference in proportions. Analogously, for a Poisson model the maximum likelihood estimate is a function of the difference in mean rates. The normal approximation to interval estimation is more erratic than stated in textbooks.[2] Let π be the probability that a randomly selected person in Puerto Rico dies before Maria. Let X denote the deaths in the 19 days before Maria and Y the deaths in the first 42 days after Maria. Since the size of population T is large, and π is not too small, each of these random variables can be modelled with a binomial distribution:[3]

$$X \sim \text{binom}(19T, \pi), \quad Y \sim \text{binom}(42T, \pi + \mu)$$

Thus, μ measures the increase in average probability of death after Maria on the 42 days following the storm's landfall. We will use $T = 3\,350\,535$, a reasonable population estimate for the middle of year 2017 based on US Census Vintage 2016 population estimates. For the baseline estimate of π, we will use the first 19 days of September 2017.

The profile likelihood method will treat π as a nuisance parameter and provide us with a way to estimate our main value of interest, $42T\mu$, through a confidence interval that will reliably result in nominal coverage. For our data, the method matched the results from the normal approximation. R code is available at bit.ly/2bTxxgp.

from 20 September to 30 September. In October, there were 3002 deaths in total.

Based on historical data (Table 1), we assume that mortality rates each year from September to October are constant. Hurricane Maria's death toll can therefore be estimated from the difference in the proportion or mean rate of deaths before and after the storm (see "Our model"). We do this by subtracting the average daily number of deaths before Maria from the average daily number of deaths after Maria, and multiplying the resulting figure by the number of days in this period:

$$\left(\frac{1317+3002}{11+31} - \frac{1582}{19}\right)(11+31)$$

This produces an estimated storm-related death toll of 822, from the date of landfall until the end of October.

It is best to incorporate statistical error through a 95% confidence interval for the difference in the number of deaths before and after Hurricane Maria. The normal approximation conditions for binomial data are satisfied, but to avoid confidence intervals that are too narrow,[1] we verified the results through the profile likelihood method (again, see "Our model"). We determined with 95% confidence that deaths due to Hurricane Maria from 20 September to 31 October were in the range (605, 1039), far above the certified death toll reported by the authorities.

Puerto Rico's population size makes any small changes in September due to births, deaths, immigration and emigration negligible. But more than 168 000 people have (at least temporarily) moved to Florida since October, and another 100 000 were slated to do so up to 31 December 2017 (nyti.ms/2zSe45A). Thus, our assessment of storm-related deaths when incorporating October data is likely to be conservative. If we were to assume that the population of the island in October was just 3.2 million (rather than 3.35 million) then our confidence interval becomes (726, 1150). Fewer residents means fewer natural deaths, hence a greater proportion of storm-related deaths in the reported mortality figures.

It is unclear if Maria is still causing deaths in Puerto Rico. In the future, when December mortality data is available from the DRO, this analysis will require adjustment to accommodate the smaller population and typical increase in mortality for that month. Nevertheless, it appears that the death toll from Hurricane Maria is far higher than the official numbers currently indicate – so much so that Governor Ricardo Rosselló has ordered a review of the data (bit.ly/2BMedLU), stating: "Every life is more than a number, and every death must have a name and vital information attached to it, as well as an accurate accounting of the facts related to their passing." ■

Roberto Rivera is an associate professor at the University of Puerto Rico-Mayagüez. His research interests are mainly environmental statistics and applied regression. He is also president of the Puerto Rican chapter of the American Statistical Association.

Wolfgang Rolke is professor of statistics at the University of Puerto Rico-Mayagüez. His research interest is the statistical analysis of data from high energy physics experiments at the Large Hadron Collider in Geneva, where he is an associate member of the CMS collaboration.

References

1. Agresti, A. and Coull, B. A. (1998) Approximate is better than exact for interval estimation of binomial proportions. *American Statistician*, **52**, 119–126.

2. Brown, L. D., Cai, T. T. and DasGupta, A. (2001) Interval estimation for a binomial proportion. *Statistical Science*, **16**, 101–133.

3. Casella, G. and Berger, R. L. (2001) *Statistical Inference* (2nd edn). Pacific Grove, CA: Duxbury.

Left: photo courtesy of Manuel Lopez

Below: photo courtesy of Ida Perez