Use Of Mathematics And Geology To Predict Possible Volcano Eruption: Personal Investigation Of The Method

A volcano is one of the world's most astonishing powers. They are explosive landforms which are able to eject molten rocks, lava, and thick ash high into the atmosphere, sometimes with devastating consequences. Volcanoes mostly occur along destructive and constructive plate boundaries where plates are pushed together or dragged apart. Cracks or weaknesses allow magma to rise up from below the earth's crust, pressure builds up which then releases suddenly, causing the magma to explode; a volcanic eruption. After more eruptions over time, the mound of rock builds up forming a cone shaped volcano which is what people typically see around the world.

The state of Hawaii is the home of many volcanoes, active or not. Seeing the mountain like shapes along the landscape of a small island in the middle of the North Pacific Ocean, made me wonder as to how people feel safe living on such a small area of space, and sharing it with active volcanoes as they are usually associated with death and destruction. In looking for an area of research I was intrigued by the idea of being able to predict where and when certain volcanoes around the world would erupt, and how this information could help people around the world and potentially save their lives. Volcanoes are thought of as natural disasters, and as seen in the past, their outcomes may be deadly.

An example is the city of Pompeii, the most famous eruption of the ancient world. Pompeii was founded in the sixth or seventh century B.C., near the Bay of Naples1. Mount Vesuvius was and still to this day is an active volcano, erupting 50 times or more to come. The famous eruption that took the entire city’s life was the eruption of the year 79 A.D., that left the city covered under volcanic ash and lava with an approximate number of 16,000 casualties from the city of Herculaneum as well2. Vesuvius is a part of the Campanian volcanic arc, which is a line of volcanoes that formed over a subduction zone created by the convergence of the African and Eurasian plates3. This zone is also the cause of many other volcanoes such as Mount Etna and Stromboli4.

I decided that I would try to predict mathematically which volcanoes are going to erupt and why, with the use of probability theory and statistical analysis, along with geography and geology. The basic principle for me was to look at the active volcanoes in the world right now, their plate tectonics and specifically the ‘Ring of Fire’, an area drowned in active volcanoes, seen in figure 1.

Figure 1. Active Volcanoes, Plate Tectonics, and the “Ring of Fire”

Along with the Poisson distribution, which is a statistical distribution that shows how many times an event, in this case being a volcanic eruption, is likely to occur within a specified period of time1.

Poisson Distribution

The Poisson Distribution is a discrete function, meaning that whichever event the math is referring to, can only be measured as occurring or not occurring1. A discrete distribution is a statistical distribution that shows the probabilities of outcomes with finite values 2 meaning the variable can only be measured out in whole numbers as fractional occurences of the event are not a part of the model3. In this case, this is ideal as volcanoes too come in whole numbers. Statistical distributions can be either discrete or continuous4. For example, when studying the probability distribution of a die with six numbered sides there can only be six possible outcomes, so the finite value is six. Another example can include flipping a coin. Flipping a coin can only result in two outcomes so the finite value is two. A continuous distribution is built from outcomes that potentially have infinite measurable values5.

I will use the Poisson distribution to estimate how likely it is that volcanic eruptions will occur, 'X' number of times. However, the poisson distribution can be used to estimate many different ranges of subject from one of the most famous historical, practical uses: the annual number of Prussian cavalry soldiers killed due to horse-kicks. Other modern examples include estimating the number of car crashes in a city of a given size; in physiology, this distribution is often used to calculate the probabilistic frequencies of different types of neurotransmitter secretions.

The Poisson distribution is:

Let X be a discrete random variable that represents the number of events observed over a given time period. Letting ? be the expected value (average) of X. If X allows a Poisson distribution, then the probability of observing k events over the time period is the equation shown above, where e is Euler's number. However, first I was curious to know how it was derived, which can be found through a problem. The problem asks to consider a binomial distribution of XB (n,p). A binomial distribution gives the discrete probability distribution of obtaining exactly n successes out of N Bernoulli trials. As P (X = k) = for k = 0, 1 , 2 , 3 , . . , n., the problems asks to take the limit of the above using n . Instead of having a p, let us assume that it is given that the mean of the probability distribution function, is value m. The problem asks to find P (X = K) in terms of m and k for a new distribution, where k = 0 , 1 , 2 , 3 , . . . ,

This formula calculates the probability that X equals a given value of k. ? is the mean of the distribution. If X represents the number of volcanic eruptions we have Pr(X ?1) = 1 – Pr(x = 0). This gives us a formula for working out the probability of an eruption as 1 -e-?. For example, as I was earlier in my introduction discussing Mount Vesuvius, I thought this would be the first volcano I look into, which is where I started running into some difficulties.

The First Problem

Volcanoes have a life cycle, however many of them do not have a specific eruption pattern that they follow. Many of them like Vesuvius, have sporadic eruption cycles. Geologists believe there were three significant Mount Vesuvius eruptions prior to the one in A.D. 79, most famously one in 1800 B.C. Until about 1631, geologists and historians believe the volcano erupted about a dozen times, with varying severity. In 1631, however, the volcano entered a period of frequent volcanic activity. From 1631 to the end of the nineteenth century, Mount Vesuvius erupted fifteen times, most violently in 1872. The most recent eruption occurred in 1944.

If we use the information of the period of frequent volcanic activity, 1900-1631 is 269 years. If

the volcano erupted 15 times during that time frame, we would assume that the volcano went off approximately every 17.933 years. However, as the period of frequency activity ended in 1900, the most recent eruption after that was 44 years later. Now it is 2019, 75 years since the latest eruption. Clearly, there is no set time frame, making it difficult to put into the equation. I could take a look at tectonic plate movements under the volcano, however that reaches too far into different subjects, not modelling volcanoes.

I decided to look at a different volcano, the Yellowstone supervolcano. It unlike Vesuvius has an eruption pattern; every 600,000 years. Therefore, ? is the number of eruptions every year, we have 1/600,000 ? 0.00000167 and 1 -e-?. ? 0.00000167. If we look at the probability over a range of years, you can modify the formula for probability as 1 -e-t? where t is the number of years for our range. Therefore, the probability of a Yellowstone eruption in the next 1000 years is 1 -e-0.00167? 0.00166, and the probability in the next 10,000 years is 1 -e-0.0167 ? 0.0164. So we have approximately a 2% chance of this eruption in the next 10,000 years. A far smaller volcano, like Katla in Iceland has erupted 16 times in the past 1100 years, giving an average eruption every ? 70 years. This gives ? = 1/70 ? 0.014. So we can expect this to erupt in the next 10 years with probability 1 -e-0.14 ? 0.0139. And in the next 30 years with probability 1 -e-0.42 ? 0.34.

I found two volcanoes that were easy to work with as they had patterns, and the eruption data was easily accessible and found. However, it led me to think about how much more complicated the process can get especially as we often don't know the accurate data for the value of ?, which is what I experienced with Mount Vesuvius. However, I came across an estimating technique called Maximum Likelihood Estimation.

Maximum Likelihood Estimation

In statistics, Maximum Likelihood Estimation is a method that determines values for the parameters of a model. The parameter values are found such that they maximise the likelihood that the process described by the model produced the data that were actually observed. I have observed 10 data points, from various volcanoes, shown in figure 3 below.

Figure 3:

Each individual dot representing, each different volcano, and the X-axis representing how long ago each volcano erupted in (10’s of years). For example, the dot to the very left between 7 and 8 or 70 and 80 years is representing Mount Vesuvius, as it had erupted 75 years ago.

First, I had to decide which model I thought best describes the process of generating the data. In this case, the data can be described as a normal distribution as from visual inspection of the graph, it suggests that a normal distribution is plausible because most of the 10 points are clustered in the middle with few points scattered to the left and the right. A normal distribution has two parameters, the mean, ?, and the standard deviation, ?. Maximum likelihood estimation is a method that will find the values of ? and ? that result in the curve that best fits the data, seen in figure 4.

Figure 4:

The 10 data points and possible normal distributions from which the data were drawn. F1 is normally distributed with mean 10 and variance 2.25 (variance is equal to the square of the standard deviation), this is also denoted f1 ? N (10, 2.25). f2 ? N (10, 9), f3 ? N (10, 0.25) and f4 ? N (8, 2.25). The goal of maximum likelihood is to find the parameter values that give the distribution that maximise the probability of observing the data. The true distribution from which the data were generated was f1 ~ N(10, 2.25), which is the blue curve in the figure above.

Conclusion

During the course of this investigation, I tried to attempt to figure out when and which specific volcanoes would erupt. My investigation was somewhat successful. Although, I figured out the probability of the Yellowstone supervolcano eruption, and the Katla in Iceland, there is room for error in my calculations.

Even though I had the data, I could not be sure that the volcano would erupt regularly because an eruption depends on so many factors and resources to understand. Further, it is generally futile to go by probability when we are faced with binary situations where 1 or 0 may ruin us completely. The question of: what is the point in trying to predict an event that may or may not happen in 600,000 years?, is very real and valid. Most people like to live in the ‘now’ and now worry about such events happening for such a long time from now, as in 600,000 many other disasters and discoveries will be found.

Another part of the probability that I had a hard time with was that the probability of Katla in Iceland, I mentioned that “we can expect this to erupt in the next 10 years with probability 1 -e-0.14 ? 0.0139. And in the next 30 years with probability 1 -e-0.42 ? 0.34.”, however there is no proof that this has to happen in the next consecutive 10 and 30 years, as it can just erupt in any given random 10/30 year range.

However, the biggest lesson I have learned from my research is that, although geologists predict mathematically when and where world disasters are going to occur, there are many factors that can affect certain events, and sometimes it is just plain random, that no math or science can calculate.