Homework 4 Problem 3

William Matzko

March 4, 2021

The ‘correct’ way to do this problem is to assume there can be a maximum of one flare in an hour, and draw from a binomial distribution to simulate the flairs with probability $p = 900/(1000 \cdot 24) = 0.0375$. However, if one uses the Poisson distribution with $\lambda = 0.0375$ and $t = 1$ (i.e. a rate parameter equal to the probability of finding one flare in an hour), then in one thousand hours it is not uncommon for two flares to occur in one hour. In ten thousand hours, it’s not terribly uncommon for 3 flares to occur in an hour. Hence, the assumption that only one flare can occur in one hour is not always valid.

Of course, one may disregard this problem and approximate the Poisson distribution with the binomial distribution, but doing this was not my first thought. The problem is about applying the Poisson distribution, so I wanted to strictly use that in the code. Hence, instead of computing the number of flares every hour, I opted to compute the number of flares *per minute*. This way, it is almost guaranteed that one will only obtain 0 or 1 flares per time bin. Technically, my code forces there to be only 0 or 1 flares per minute. However, you can easily check that this is a very solid assumption:

```
import numpy as np
from collections import Counter
Counter(np.random.poisson(p, 10000))
```

The above code will count the number of flares per time bin (i.e. how many flares occur per hour in 10000 hours, or how many flares occur per minute in 10000 minutes, depending on how p is specified). Using $p = 0.0375$, we see that 2 flares (or even 3) can occur per hour. However, using $p = 0.0375/60$, we are almost guaranteed to get only 0 or 1 flares per minute (I haven’t encountered a scenario where I got more than 1 flare per minute).

I believe this code produces, more or less, the same result as the ‘correct’ way of doing this problem (judging by the plots Jim sent me for comparison). While this method is certainly more roundabout, and perhaps unnecessary, it was my first thought on how to tackle the problem. Since it appears to give reasonably correct results, I’ll stick with it for my final answer.

I agree that your approach is better in the sense that it avoids the violation of the “one event per dt” assumption of the Poisson. What your result tells me is that this this assumption does not explain the difference between the data and Poisson and Binomial. I posted some comments on the HW page with some thoughts on the explanation.