

P1. Study the example “How to order Reddit submissions” from Davidson-Pilon(2016b) chapter 4.

P2. Add to the Burglary system in PYMC the following situations and recompute the Burglary probability:

- a) A flooding situation;**
- b) A flooding situation and an animal permission (in a particular space).**

P3. Replacing German tanks with locomotives bring on a similar problem: a railroad numbers its locomotives in order 1..N and the locomotives with the numbers 10, 256, 202 and 97 were observed, estimate how many locomotives the railroad has. But this time you have more background information. It is probably not reasonable to assume that a train-operating company with 1000 locomotives is just as likely as a company with only 1 (as we assumed in the uniform prior for N). In most fields, there are many small companies, fewer medium-sized companies, and only one or two very large companies. In fact, the distribution of company sizes tends to follow a power law. This law suggests that if there are 1000 companies with fewer than 10 locomotives, there might be 100 companies with 100 locomotives, 10 companies with 1000, and possibly one company with 10,000 locomotives. Mathematically, a power law means that the number of companies with a given size is inversely proportional to size, or

$$\text{PMF}(x) \propto \left(\frac{1}{x}\right)^{\alpha}$$

where PMF(x) is the probability mass function of x and α is a parameter that is often near 1

- a. Estimate N in this situation**
- b. What if there are many companies with different numbers of trains, and you are equally likely to see any train operated by any company? Note that, in this case, the likelihood function is different because you are more likely to see a train operated by a large company?**