Worksheet 17

1. A fridge manufacturer tests how long it takes before their devices need to be repaired. Looking at 4 randomly selected units, it took 450, 1200, 1500, and 2400 days until replacement. (a) describe the distribution in terms of one of our known probability densities and (b) estimate the parameters of the model.

2. You want to estimate what proportion of cities in the US have over 1 million people in them. All you know is that the average city has an average 20k people. Construct a probability model for this question and make an estimate of the proportion of cities with more than 1 million people. Hints: Do this question in 'thousands of people'. Also, remember that the Pareto distribution has a mean $\frac{\alpha}{\alpha-1}$ and density:

$$f(x) = \frac{\alpha}{x^{\alpha+1}}, x > 1.$$

3. There are only 10 US cities with more than 1 million people. Ways of counting the number of cities differ, but there are somewhere between 3000 and 20000 cities in the US. How does your estimate from question 2 compare to the actual data?

4. During my last four office hours I had: 1, 2, 2, and 4 students.[1] Build a probability model and estimate the parameter(s) that describe it given the data. Predict the probability that nobody shows up to my next office hours.

[1] I made this data up. Almost nobody comes to my office hours. :(

5. A school group is giving out T-Shirts in the dining hall. Each shirt has a tag with an number on it; the first shirt was numbered 1, the second was 2, the third 3, and so forth. Shirts were handed out at random, regardless of the number. When you arrive (late) for lunch, only five shirts remain. They have the numbers 10, 45, 102, 210, and 230. Build a probability model the describes the situation (approximately). Estimate the total number of T-Shirts that were handed out using both the (a) MLE and (b) the MME.[2]

[2] This general problem is called the German Tank Problem due to its application during WWII. German tanks had a serial number on them and the Allies wanted to estimate the total number of tanks given the serial numbers from those tanks that they captured.