1. Despite seeming like the rarer phenomenon, earthquakes actually kill more people than hurricanes do.

   True. This is due in part because no one has any warning before an earthquake hits, whereas nowadays they have good information about when a hurricane will hit.

2. Forecasting earthquakes is waaaaaay easier than predicting earthquakes.

   True. "Predicting" is a definitive and specific statement about when and where an earthquake will strike. "Forecasting" is a probabilistic statement over a longer time period: there's a 60 percent chance of an earthquake in L.A. over the next thirty years.

3. If you compare how often earthquakes of various magnitudes occur worldwide, what relationship will you find?

   The number of earthquakes of a given magnitude drops off exponentially as the magnitude increases.

4. What kind of statistical distribution has the highly useful property that you can estimate the number of large-scale events from the number of small-scale ones, or vice versa?

   A power-law. (a.k.a. "scale-free.", a.k.a. "fat-tailed.")

5. In systems with noisy data and underdeveloped theory – like earthquake science and much of economics and political science – people start to mistake the **noise** for a **signal**.

6. Suppose we compare a magnitude-7.0 earthquake (on the moment magnitude scale, not the Richter), compared to a magnitude-5.0 earthquake. Which is true?

   ✓ The 7.0 quake is orders of magnitude more powerful than the 5.0 quake
   ☐ The 7.0 quake is four times as powerful as the 5.0 quake
   ☐ The 7.0 quake is twice as powerful as the 5.0 quake
   ☐ The 7.0 quake has about the same power as the 5.0 quake
   ☐ The 7.0 quake is half as powerful as the 5.0 quake
   ☐ The 7.0 quake is one-fourth as powerful as the 5.0 quake
   ☐ The 7.0 quake is orders of magnitude less powerful than the 5.0 quake

   It's 32^2, or 1024 times as powerful, to be precise.

7. What makes a model "overfit," and what problem does that cause?

   A model that is too hyper-tuned to the small number of data points it was trained on, and which therefore makes poorer predictions when given new data.