This is an optimization problem and we can apply a variety of optimization techniques. Essentially we can choose a starting point, then explore the space to see if we are getting closer to an optimum (hill-climbing, linear programming, etc). In Stan, the `optimize` function uses L-BFGS.

Probability Density Functions

A PDF is a function that for a given probability distribution defines the expected histogram for long-run outcomes. Here's the PDF graphically for the Normal:

And the corresponding formula.

What we might need to do is figure out the precise shape of this curve, given some input data. Pretend for a second that you only have data, like in this figure:

You might be tempted to say "there's no pattern here" and throw up your hands. You would annoy statisticians with this approach, however. It turns out a lot of problems that generate data can be approximated (modeled) using probability distrubtions. FOr example, the arrival rate of cars at a red light can be modeled with the Poisson distribution.