**Bootstrap PDF’s**

Say we have some data,

|  |
| --- |
| Weight |
| 27 |
| 31 |
| 26 |
| 25 |
| 29 |
| 33 |

etc. We can plot a histogram of the data to get a sense of the probability distribution of the weight. Another slightly different way to get the histogram is to do repeated sampling of the data, w/r replacement. Can do this is N times, say, and this is called a bootstrap sample I think. Then we can make a histogram of all bootstrap sample. This should look the same of course, for large sample size, N, because our bootstrap sample should just end up being a bunch of copies (N/6, roughly) of our original data.

We can extend this bootstrapping technique to other scenarios, however. Say we wanted the approximate probability distribution of the mean, weight, = (1/n)Σi=1nWi (where n would have to be much less than the length of our data set). Then we could create a bootstrap sample of by picking n of the guys out (w/ replacement), finding their average, and putting that average in our bootstrap sample. We’d do this N times, and then create a histogram. And we can do this for more than just the average. We could construct a histogram of the variance this way too. Or of more exotic things like the median, or mode. Thus, while we may not be able to mathematically specify what the sampling distribution of some statistic is, we can often plot it at least, via the bootstrap method.