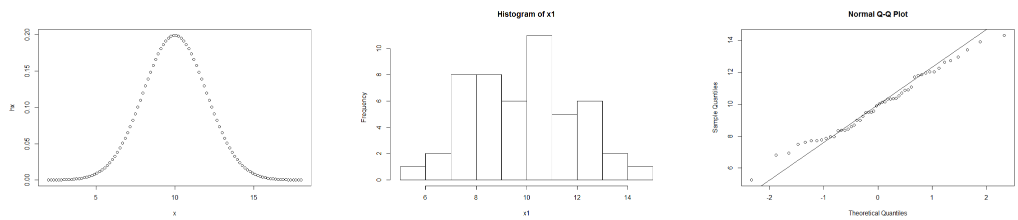
Math 216
§3.2 Probability Plots
More Examples

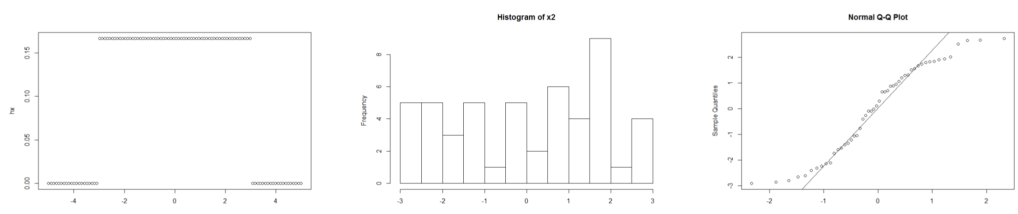
In each of the examples below, the probability density function on the left was used to generate 50 random observations, summarized in the accompanying histogram. The observations were then used to construct the normal probability plot on the right. Note that in each case, the y-values of the points on the probability plot are the values of the individual observations.

Example 1:



In this example, the 50 data points came from a normal random variable with mean 10 and standard deviation 2. The resulting probability plot is very linear. A few points at the low and high ends of the range set lie off the line. That's because even 50 data points isn't enough to represent the tails of the normal distribution very well.

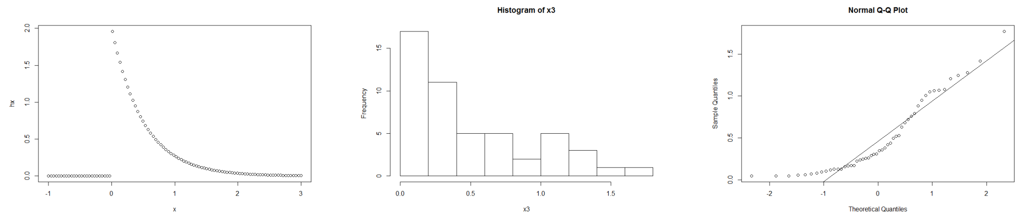
Example 2:



This time the data were generated from a uniform distribution in which every value between -3 and 3 is equally likely. Note that this interval is represented on the vertical axis of the probability plot—all the y-coordinates of the points on the scatterplot lie between -3 and 3.

This time the data at the low end of the range are well above where they would be if they were normally distributed. This indicates that the distribution has a much shorter left-hand tail than the normal distribution, which is true of the uniform distribution. Likewise, the data at the high end of the range are well below where they would be if they were normally distributed, indicating a short right-hand tail to the distribution.

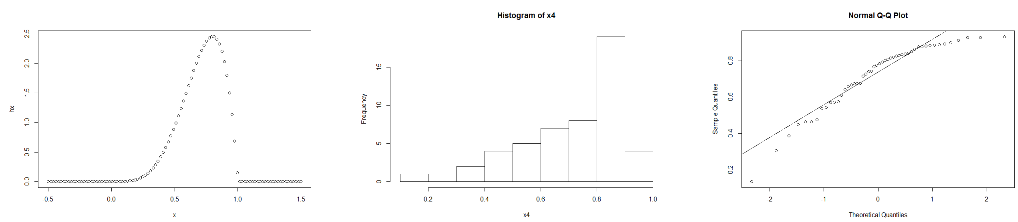
Example 3:



The distribution here is an exponential distribution, which has a relatively long right-hand tail compared to the normal distribution. We can see that in the probability plot: the data at the high end of the range are just a bit larger than we would expect for normally distributed data. With an exponential distribution, we get large outliers a little more frequently than with a normal distribution.

The exponential distribution doesn't allow any negative data, so the left-hand tail is very short. We can see that in the probability plot: the data at the low end of the range are much larger than we would expect for normal data, and they're all positive.

Example 4:



This is a beta distribution. It's a bit smoother than the exponential distribution in Example 3, but it's also strongly skewed. Where the exponential distribution was skewed to the right, the beta is skewed to the left. As a result, the probability plot here is just the opposite of the one in Example 3: this is one is concave down where that one was concave up.

The beta distribution has a long left-hand tail, so we get a few more small outliers than we would with a normal distribution. We see those outliers sitting well below the line in the probability plot. The beta distribution has a very short right-hand tail, and, in fact, the values generated by this beta distribution never exceed 1. That gives us something of an asymptote in the probability plot at $y = 1$.