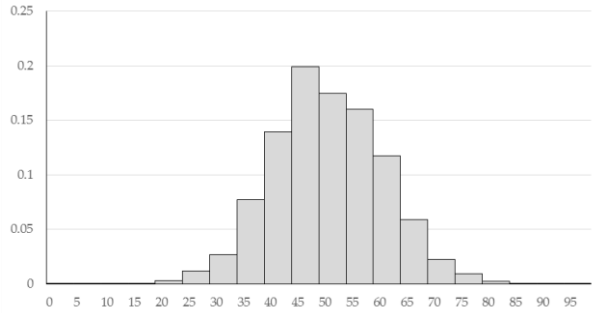
Part 9: Looking at Graphs

We are sometimes given graphs and ask to choose what distribution best fits. There are a few things that we can do to check. Below we will look at a few examples.

Example (Normal)



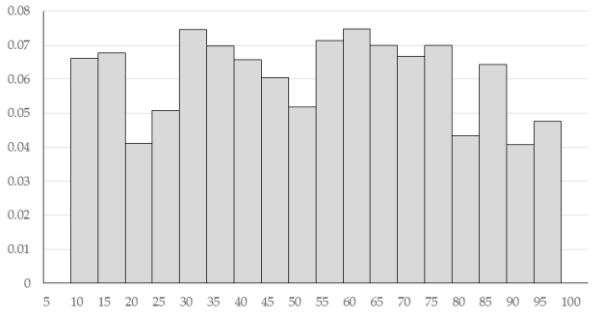
The normal distribution is a continuous distribution and therefore should be represented by a histogram. It should be reasonably even on both sides of the centre, i.e. symmetrical, and generally follow a bell shape with most of the data in the centre and less data towards the edge.

The two parameters we estimate are the mean and the standard deviation.

To estimate the mean look for where the centre of the data is, and if the bars are taller to the right or the left move your mean just slightly in that direction. For example, the graph above has a mean of approximately 48.

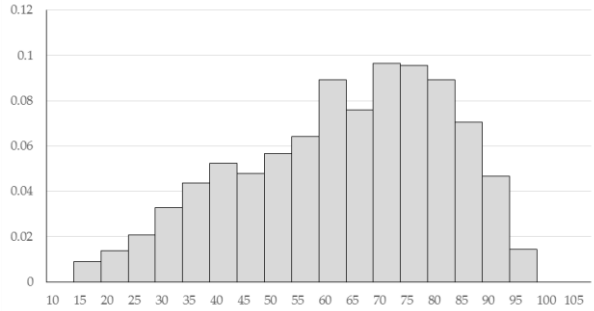
To estimate the standard deviation look for the smallest data point (in this case 20) and the largest data point (in this cast 80, as the data between 80 and 85 is too small to count) and then divide by 6. So in this case I would estimate the standard deviation to be 10 (60 ÷ 6).

Example (Rectangular)

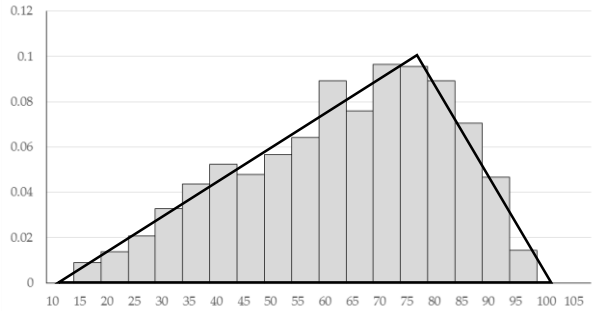


The rectangular distribution is also a continuous distribution and therefore should be represented by a histogram. It should be relatively even along, and then just suddenly stop. The two parameters that are needed are the minimum (in this case 10) and the maximum (in this case 100)

Example (Triangular)

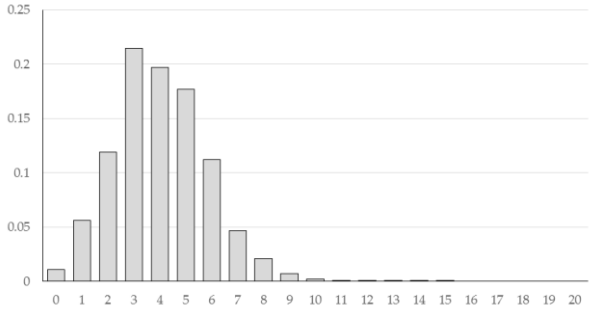


The triangular distribution is also a continuous distribution and therefore should be represented by a histogram. The triangular distribution is most useful when the data is skewed in one direction or the other and we need to work out three parameters, the minimum, the maximum and the most likely. Often the easiest way to do this is to draw a triangle over the data to see where these points occur. For example:



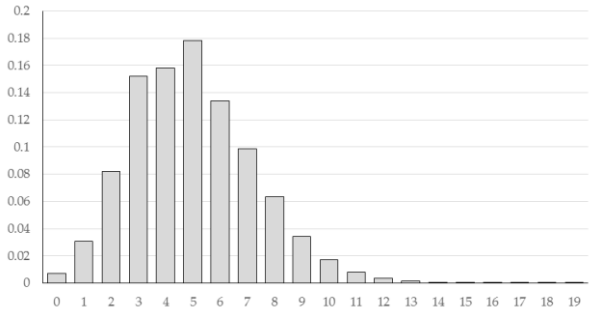
You want the lines of the triangle to evenly cut the open space and the bars as shown above. This would give us a minimum value of 12 (a), a maximum value of 102 (b) and a most likely value of 78 (c).

Example (Binomial)



The binomial distribution is a discrete distribution and therefore it should be represented by a bar graph (rather than a histogram).

Example (Poisson)



The Poisson distribution is a discrete distribution and therefore it should also be represented by a bar graph (rather than a histogram).

As you may have noticed these two graphs are very similar to each other. The big difference between a binomial distribution and a Poisson distribution is a binomial has a fixed number of trials, so will therefore stop after a certain point, in the graph above you can see that there is no more data after 15, so there must be 15 trials, however the easiest way to work what type of distribution it is will be to look at the **context** of the question that will be given with it.