**Key points**

* Many common data science tasks can be framed as estimating a parameter from a sample.
* We illustrate statistical inference by walking through the process to estimate p. From the estimate of p, we can easily calculate an estimate of the spread, 2p−1.
* Consider the random variable X that is 1 if a blue bead is chosen and 0 if a red bead is chosen. The proportion of blue beads in N draws is the average of the draws X1,...,XN.
* X¯ is the *sample average*. In statistics, a bar on top of a symbol denotes the average. X¯ is a random variable because it is the average of random draws - each time we take a sample, X¯ is different.

X¯=X1+X2+...+XNN

* The number of blue beads drawn in N draws, NX¯, is N times the proportion of values in the urn. However, we do not know the true proportion: we are trying to estimate this parameter p.