# Estimate of population and sub-population sizes

The size of *D* and *M* are fundamental properties of the hypergeometric stochastic system like *p* for a binomial process and *l* for a Poisson process. Distributions of our uncertainty about the value of these parameters can be determined from Bayesian inference, given a certain sample size taken from the population *M*, of which s belonged to the sub-population *D*. The hypergeometric probability of *s* successes in *n* samples from a population *M* of which *D* have the characteristic of interest is given as:

| | | | |
|---|---|---|---|
| $p(s) = \frac{\binom{D}{s}\binom{M-D}{n-s}}{\binom{M}{n}}$ | $0 \le s \le n\,,$ | $s \le D\,,$ | $n \le M$ |

So, with a Uniform prior, we get the following posterior densities for *D* and *M*:

$$p(D) = \frac{\binom{D}{s}\binom{M-D}{n-s}}{\binom{M}{n}} \propto \frac{D!(M-D)!}{(D-s)!(M-D-n+s)!}$$

$$p(M) = \frac{\binom{D}{s}\binom{M-D}{n-s}}{\binom{M}{n}} \propto \frac{(M-D)!(M-n)!}{(M-D-n+s)!M!}$$

These formulae do not equate to standard distributions, and need to be normalized, which can be done easily with Crystal Ball and Excel as shown in the Excel example in the previous topic.