# Estimation of the number of trials n made after having observed s successes with probability p

### The problem

Consider the situation where we have observed *s* successes and know the probability of success *p*, but would like to know how many trials were actually done to have observed those successes. We wish to estimate a value that is fixed, so we require a distribution that represents our *uncertainty* about what the true value is. There are two possible situations: we either know that the trials stopped on the $s^{th}$ success or we do not.

### Results

If we know that the trials stopped on the $s^{th}$ success, we can model our uncertainty about the true value of *n* as:

*n* = NegBinomial(p,*s*)

If, on the other hand, we do not know that the last trial was a success (though it could have been), then our uncertainty about *n* is modeled as:

*n* = NegBinomial(p, *s*+1) - 1

Both of these formulae result from a Bayesian analysis with Uniform priors for *n*.

### Derivations

Let *x* be the number of trials that were needed to obtain the $s^{th}$ success. We will use a uniform prior for *x*, i.e. *p(x)* = *c*, and, from the binomial distribution, the likelihood function is the probability that at the $(x\text{-}1)^{th}$ trial there had been (*s*-1) successes *and* then the $x^{th}$ trial was a success, which is just the Negative Binomial probability mass function:

$$l(X \mid x) = \binom{x-1}{x-s} p^s (1-p)^{x-s}$$

Since we are using a uniform prior (assuming no prior knowledge), and the equation for *l(X|x)* comes directly from a distribution (so it must sum to unity) we can dispense with the formality of normalizing the posterior distribution to one, and observe:

$$p(x) = \binom{x-1}{x-s} p^s (1-p)^{x-s}$$

i.e. that *x* = NegBinomial(p,*s*).

In the second case, we do *not* know that the last trial was a success, only that in however many trials were completed, there were just *s* successes. We have the same Uniform prior for the number of trials, but our likelihood function is just the binomial probability mass function, i.e.:

$$l(X \mid x) = \binom{x}{s} p^s (1-p)^{x-s}$$

Since this does not have the form of a probability mass function of a known distribution, we need to complete the Bayesian analysis, so:

$$f(x \mid X) = \frac{\binom{x}{s} p^s (1-p)^{x-s}}{\sum_{i=s}^{\infty} \binom{i}{s} p^s (1-p)^{i-s}}$$

Look at the denominator, and substituting *j = i+1* gives:

$$\sum_{i=s}^{\infty} \binom{i}{s} p^s (1-p)^{i-s} = \sum_{j=s+1}^{\infty} \binom{j-1}{s} p^s (1-p)^{j-s-1} = \left(\frac{1}{p}\right). \sum_{j=s+1}^{\infty} \binom{j-1}{s} p^{s+1} (1-p)^{j-s-1} = \left(\frac{1}{p}\right)$$

since

$$\binom{j-1}{s} p^{s+1} (1-p)^{j-s-1}$$

is the probability mass function for the NegBinomial(s+1,p) distribution for j and therefore sums to 1. The posterior distribution then reduces to:

$$f(x \mid X) = \binom{x}{s} p^{s+1} (1-p)^{x-s}$$

For x = y - 1:

$$f(y \mid X) = \binom{y-1}{s} p^{s+1} (1-p)^{y-s-1}$$

i.e. y = NegBinomial(p,s+1), and therefore x = NegBinomial(p,*s*+1) -1 distribution.