1 A Model of the Negative Binomial Distribution

The model is based on the following experiment:

Consider a coin for which the probability of flipping a head (denoted H) is p. Then the probability of flipping a tail (denoted T) is $q = 1 - p$. The experiment consists of repeatedly tossing the coin until one has flipped n heads. Once this has been accomplished one stops. The sample space Ω for this problem is comprised of words built out of the letters H and T and which contain H exactly n times and which contain T ℓ times, where ℓ could take on any value $\{0, 1, 2, 3, \ldots\}$. Let X denote the discrete random variable defined as

$$X(\omega) = \text{ the number of tosses in outcome } \omega.$$

We would like to determine the probability of the event that it takes exactly k tosses to get n heads; i.e. $P(X = k)$. (Because we must get at least n heads, k must be greater than or equal to n.) Clearly, one has the relation

$$k = n + \ell$$

since the total number of tosses, k, must equal the sum of the number of heads tossed (n) plus the number of tails tossed (ℓ).

The number of outcomes corresponding to the event of flipping exactly n heads in k successive tosses is clearly equal to the number of words that can be made from k letters, which are either H or T, if the last letter must be an H. This is equal to the number of words that can be made from $k - 1$ letters of which $n - 1$ are H and all the others are T. By Theorem 4 of *Notes on Counting* this number is

$$\binom{k-1}{n-1} = \binom{n+\ell-1}{\ell}. \tag{1}$$

It follows from (1) that

$$P(X = k) = \binom{k-1}{n-1} p^n q^{k-n} = \binom{n+\ell-1}{\ell} p^n q^{\ell}. \tag{2}$$

In order to show that this actually is a probability mass function, we must show that

$$\sum_{k=n}^{\infty} P(X = k) = 1.$$

Showing that is equivalent to the following calculation:

$$\sum_{k=n}^{\infty} P(X = k) = \sum_{k=n}^{\infty} \binom{k-1}{n-1} p^n q^{k-n} \tag{3}$$

$$
\begin{aligned}
&= \sum_{\ell=0}^{\infty} \binom{n+\ell-1}{\ell} p^n q^\ell && (4)\\
&= p^n \sum_{\ell=0}^{\infty} \binom{n+\ell-1}{\ell} q^\ell && (5)\\
&= p^n (1-q)^{-n} = p^n p^{-n} = 1 && (6)
\end{aligned}
$$

To understand how one gets from equation (5) to equation (6) above, rewrite $(1-q)^{-n}$ as

$$
\begin{aligned}
(1-q)^{-n} &= \left(\sum_{\ell_1=0}^{\infty} q^{\ell_1}\right) \cdots \left(\sum_{\ell_n=0}^{\infty} q^{\ell_n}\right) && (7)\\
&= \sum_{\ell=0}^{\infty} \alpha_\ell q^\ell. && (8)
\end{aligned}
$$

The coefficient α_ℓ equals the number of ways there are to choose ℓ objects from n different types of objects when there are an unlimited number of each type of object. The *different types of objects* here correspond to the terms in the different geometric series. More precisely, q^{ℓ_i} corresponds to choosing ℓ_i objects of type i. A particular choice of ℓ objects from the various types corresponds to one term in the product of the series which has the form q^ℓ where $\ell = \ell_1 + \ldots + \ell_n$. We are thus looking at the number of ways there are to write ℓ as $\ell = \ell_1 + \ldots + \ell_n$. The answer to this counting problem is, by Theorem 7 of *Notes on Counting*,

$$\alpha_\ell = \binom{n+\ell-1}{\ell}.$$

This justifies the transition from (5) to (6).

Note One could have made this last argument to justify the right hand side of (1) directly; i.e., the event of tossing a coin k times until a the n^{th} head turns up is equivalent to counting the outcomes of the form ℓ_1 tails followed by a head and then ℓ_2 tails followed by a head, etc. up to ℓ_n tails followed by an n^{th} final head, where $\ell = \ell_1 + \ldots + \ell_n$ and $k = n + \ell$.