Binomial and Multinomial Coefficients

The **binomial coefficient** allows one to compute the number of combinations of *N* things taken *n* at a time. The order is not important and no repetitions are permitted. Think of *N* as the number of weasels in a defined population and let *n* be the sample size. The binomial coefficient can be used to compute the number of possible ways a sample of size *n* can be taken from the population of *N* individuals. The binomial coefficient is written as

$$\binom{N}{n} = \frac{N!}{n!(N-n)!} ,$$

where the symbol ! means factorial (e.g., 5! = 5x4x3x2x1 = 120). The notation $\binom{N}{n}$ is usually read "*N* choose *n*."

Now, how many ways can a sample of size 2 be taken from a population of size 4 weasels (i.e., 4 choose 2")?

$$\binom{4}{2} = \frac{4!}{2!(4\text{-}2)!} = \frac{4x3x2x1}{2x1\ (2x1)} = 6.$$

This is a simple example, so we can check it easily by enumerating all the possible combinations. Let the 4 "things" be A, B, C, and D. How many ways can two of these be taken from a total of 4? We have

AB, AC, AD, BC, BD, and CD

a total of 6 ways. This agrees with the computation using the binomial coefficient (above). Note, AB and BA are the same here; they do not count twice. The binomial coefficient is a counting shortcut.

This coefficient is handy when you have a population of 100 salamanders (*N*) and need to know the number of samples of size 30 that are possible (e.g., *N* = 100 and *n* = 30). The answer is about 2.9 trillion trillion (program *DERIVE* will compute this in about 10 seconds using its command COMB(100,30)). Clearly, one needs a handy shortcut, as the methodical counting of all such combinations is nearly always impossible in practical problems.

The idea here is useful in understanding expected values of discrete random variables and in random sampling.

The **multinomial coefficient** is an extension of the binomial coefficient and is also very useful in models developed in fw663. The multinomial coefficient is nearly always introduced by way of die tossing.

$$\binom{N}{n_1\ n_2\ n_3\ n_4\ n_5\ n_6} = \frac{N!}{n_1!\,n_2!\,n_3!\,n_4!\,n_5!\,n_6!}$$

$$= \frac{N!}{\prod n_i!}\ .$$

Note the use of the product operator $\prod$ in the last expression; it is similar to the summation operator $\sum$. It is handy in many instances in statistics.

A property of multinomial data is that there is a dependency among the counts of the 6 faces. For example, if a die is thrown and it is not a 1, 3, 4, 5, or 6, then it *must* be a 2. One can always make this dependency clear, such as writing the final term in the denominator as

$$n_6 = N - \sum_{i=1}^{5} n_i\ .$$

While this may appear clumsy, we will see much of this type of notation. Note that the binomial has this same dependency; if we know *n* and the number of heads, we then know the number of tails. Knowing any two of the three statistics {*n*, *y* or (*n*-*y*)}, allows knowledge of the third.