Properties of Hypergeometric Distribution

The Hypergeometric distribution is given by

$$P_r = \frac{\binom{w}{r}\binom{n-w}{m-r}}{\binom{n}{m}}.$$

Firstly, we will show that

$$\sum_{r=0}^{\min(w,m)} P_r = 1.$$

Using Vandermonde's identity we find

$$\sum_{r=0}^{\min(w,m)} \binom{w}{r}\binom{n-w}{m-r} = \binom{n}{m}.$$

Thus it is clear that

$$\sum_{r=0}^{\min(w,m)} P_r = 1.$$

Now we will evaluate the expectation value $\langle r \rangle$

$$\langle r \rangle = \sum_{r=0}^{\infty} rP_r = \sum_{r=0}^{\min(w,m)} rP_r = \frac{\sum_{r=0}^{\min(w,m)} r\binom{w}{r}\binom{n-w}{m-r}}{\binom{n}{m}}$$

Now we make use of the following

$$\binom{n}{m} = \frac{n}{m}\binom{n-1}{m-1}.$$

Thus

$$\begin{aligned}
\frac{\sum_{r=0}^{\min(w,m)} r\binom{w}{r}\binom{n-w}{m-r}}{\binom{n}{m}} &= \frac{\sum_{r=0}^{\min(w,m)} r\frac{w}{r}\binom{w-1}{r-1}\binom{n-w}{m-r}}{\binom{n}{m}} \\
&= w\frac{\sum_{r=0}^{\min(w,m)} \binom{w-1}{r-1}\binom{n-w}{m-r}}{\binom{n}{m}} \\
&= w\frac{\binom{n-1}{m-1}}{\binom{n}{m}} = w\frac{\binom{n}{m}\frac{m}{n}}{\binom{n}{m}} = w\frac{m}{n}
\end{aligned}$$

Now we consider the variance $\sigma^2 = \langle r^2 \rangle - \langle r \rangle^2$. Here we will evaluate $\sigma^2 = \langle r(r-1) \rangle + \langle r \rangle - \langle r \rangle^2$. Evaluating the first term,

$$\langle r(r-1) \rangle = \sum \frac{r(r-1)}{\begin{pmatrix} n \\ m \end{pmatrix}} \left(\frac{w}{r}\right) \left(\frac{w-1}{r-1}\right) \begin{pmatrix} w-2 \\ r-2 \end{pmatrix} \begin{pmatrix} n-w \\ m-r \end{pmatrix} = \frac{w(w-1)}{\begin{pmatrix} n \\ m \end{pmatrix}} \begin{pmatrix} n-2 \\ m-2 \end{pmatrix}$$

where we have applied Vandermonde's identity. Some simplification gives,

$$\langle r(r-1) \rangle = \frac{w(w-1)}{\begin{pmatrix} n \\ m \end{pmatrix}} \begin{pmatrix} n \\ m \end{pmatrix} \frac{m(m-1)}{n(n-1)} = \frac{w(w-1)m(m-1)}{n(n-1)}.$$

Now recall we have

$$\langle r \rangle = \frac{wm}{n}$$

and hence

$$\langle r \rangle^2 = \frac{w^2 m^2}{n^2}.$$

Putting these into

$$\sigma^2 = \langle r(r-1) \rangle + \langle r \rangle - \langle r \rangle^2$$

after a little more algebra we obtain,

$$\sigma^2 = m \frac{w}{n} \left(1 - \frac{w}{n}\right) \frac{(n-m)}{(n-1)}.$$