proof of expected value of the hypergeometric distribution

We will first prove a useful property of binomial coefficients. We know

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}.$$

This can be transformed to

$$\binom{n}{k} = \frac{n}{k}\frac{(n-1)!}{(k-1)!(n-1-(k-1))!} = \frac{n}{k}\binom{n-1}{k-1}. \tag{1}$$

Now we can start with the definition of the expected value:

$$E[X] = \sum_{x=0}^{n} \frac{x\binom{K}{x}\binom{M-K}{n-x}}{\binom{M}{n}}.$$

Since for $x = 0$ we add a 0 in this formula we can say

$$E[X] = \sum_{x=1}^{n} \frac{x\binom{K}{x}\binom{M-K}{n-x}}{\binom{M}{n}}.$$

Applying equation (**??**) we get:

$$E[X] = \frac{nK}{M}\sum_{x=1}^{n} \frac{\binom{K-1}{x-1}\binom{M-1-(K-1)}{n-1-(x-1)}}{\binom{M-1}{n-1}}.$$

Setting $l := x - 1$ we get:

$$E[X] = \frac{nK}{M}\sum_{l=0}^{n-1} \frac{\binom{K-1}{l}\binom{M-1-(K-1)}{n-1-l}}{\binom{M-1}{n-1}}.$$

The sum in this equation is 1 as it is the sum over all probabilities of a hypergeometric distribution. Therefore we have

$$E[X] = \frac{nK}{M}.$$

proof of variance of the hypergeometric distribution*

We will first prove a useful property of binomial coefficients. We know

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}.$$

This can be transformed to

$$\binom{n}{k} = \frac{n}{k}\frac{(n-1)!}{(k-1)!(n-1-(k-1))!} = \frac{n}{k}\binom{n-1}{k-1}. \tag{1}$$

The variance $\mathrm{Var}[X]$ of X is given by:

$$\mathrm{Var}[X] = \sum_{x=0}^{n}\left(x - \frac{nK}{M}\right)^2 \frac{\binom{K}{x}\binom{M-K}{n-x}}{\binom{M}{n}}.$$

We expand the right hand side:

$$\begin{aligned}\mathrm{Var}[X] \quad = \quad & \sum_{x=0}^{n}\frac{x^2\binom{K}{x}\binom{M-K}{n-x}}{\binom{M}{n}} \\ & -\frac{2nK}{M}\sum_{x=0}^{n}\frac{x\binom{K}{x}\binom{M-K}{n-x}}{\binom{M}{n}} \\ & +\frac{n^2K^2}{M^2}\sum_{x=0}^{n}\frac{\binom{K}{x}\binom{M-K}{n-x}}{\binom{M}{n}}.\end{aligned}$$

The second of these sums is the expected value of the hypergeometric distribution, the third sum is 1 as it sums up all probabilities in the distribution. So we have:

$$\mathrm{Var}[X] = -\frac{n^2K^2}{M^2} + \sum_{x=0}^{n}\frac{x^2\binom{K}{x}\binom{M-K}{n-x}}{\binom{M}{n}}.$$

In the last sum for $x=0$ we add nothing so we can write:

$$\mathrm{Var}[X] = -\frac{n^2K^2}{M^2} + \sum_{x=1}^{n} \frac{x^2\binom{K}{x}\binom{M-K}{n-x}}{\binom{M}{n}}.$$

Applying equation (**??**) and $x = (x-1)+1$ we get:

$$\mathrm{Var}[X] = -\frac{n^2K^2}{M^2} + \frac{nK}{M}\sum_{x=1}^{n} \frac{(x-1)\binom{K-1}{x-1}\binom{M-K}{n-x}}{\binom{M-1}{n-1}} + \frac{nK}{M}\sum_{x=1}^{n} \frac{\binom{K-1}{x-1}\binom{M-K}{n-x}}{\binom{M-1}{n-1}}.$$

Setting $l := x-1$ the first sum is the expected value of a hypergeometric distribution and is therefore given as $\frac{(n-1)(K-1)}{M-1}$. The second sum is the sum over all the probabilities of a hypergeometric distribution and is therefore equal to 1. So we get:

$$\begin{aligned}
\mathrm{Var}[X] &= -\frac{n^2K^2}{M^2} + \frac{nK(n-1)(K-1)}{M(M-1)} + \frac{nK}{M} \\
&= \frac{-n^2K^2(M-1) + Mn(n-1)K(K-1) + KnM(M-1)}{M^2(M-1)} \\
&= \frac{nK(M^2 + (-K-n)M + nK)}{M^2(M-1)} \\
&= \frac{nK(M-K)(M-n)}{M^2(M-1)} \\
&= n\frac{K}{M}\left(1-\frac{K}{M}\right)\frac{M-n}{M-1}.
\end{aligned}$$

This formula is the one we wanted to prove.