Statistical Inference Chapter 3

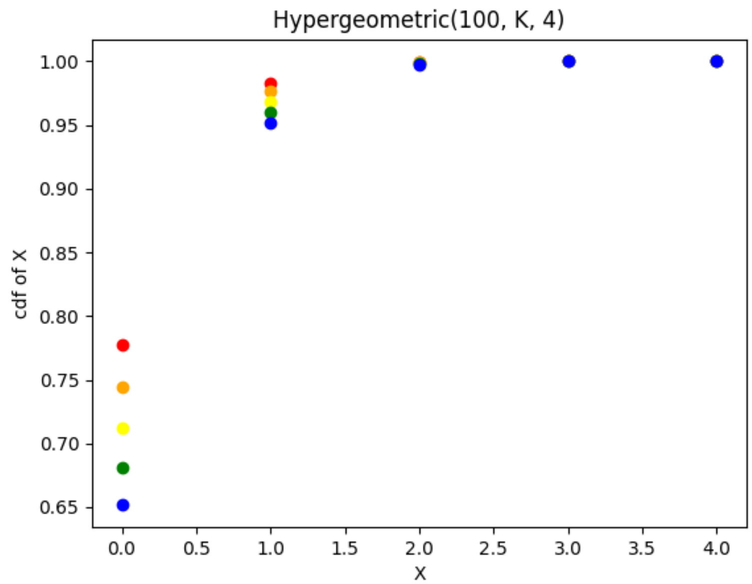
Gallant Tsao

July 15, 2024

1. We first note that the pmf of X is
$$p_X(x) = \frac{1}{N_1 - N_0 + 1}, \ x \in \{N_0, N_0 + 1, ..., N_1\}.$$
Then we get the expectation to be
$$\begin{aligned}\mathbb{E}[X] &= \sum_{x=N_0}^{N_1} x \frac{1}{N_1 - N_0 + 1} \\ &= \frac{1}{N_1 - N_0 + 1} \cdot \frac{N_1 - N_0 + 1}{2}(2N_0 + (N_1 - N_0 + 1 - 1)) \\ &= \frac{N_1 + N_0}{2}.\end{aligned}$$
As for the variance, we get
$$\begin{aligned}\mathbb{E}[X^2] &= \sum_{x=N_0}^{N_1} x^2 \frac{1}{N_1 - N_0 + 1} \\ &= \frac{1}{N_1 - N_0 + 1}\Big(\sum_{x=1}^{N_1} x^2 - \sum_{x=1}^{N_0 - 1} x^2\Big) \\ &= \frac{1}{N_1 - N_0 + 1}\Big(\frac{N_1(N_1+1)(N_1+2) - (N_0 - 1)(N_0)(2N_0 - 1)}{6}\Big)\end{aligned}$$
So that
$$\begin{aligned}Var(X) &= \mathbb{E}[X^2] - (\mathbb{E}[X])^2 \\ &= 1\end{aligned}$$

2. Let X = number of defective parts in the sample. Then $X \sim \text{Hypergeometric}(100, n, K)$.
   (a) Firstly, we need $n = 6$ because for the same K, increasing n decreases the value of the Hypergeometric pmf (image shown at end of answer). Then with $n = 6$,
$$\begin{aligned}P(X = 0 | n = 6) &= \frac{\binom{6}{0}\binom{94}{K}}{\binom{100}{K}} \\ &= \frac{(100-k)\cdots(100-K-5)}{100 \cdots 95}\end{aligned}$$

After some trial and error with the calculations, we have that when $K = 31$, $P(X = 0) = 0.10056$, but when $K = 32$, $P(X = 0) = 0.09182$. Therefore, the sample size must be at least 32.



(b) By the same reasoning above, we need $n = 6$. Then