

One more fact about hypergeometric random variables.

If  $N$  is large and  $n$  is small, the fact that we do not replace items after they are picked because almost irrelevant, since the population  $N$  is large and we are just picking a few,  $n$ , so it doesn't matter too much if they are replaced.

Such a hypergeometric is very similar in distribution to a Binomial random variable, with same  $n$  picks and  $p = \frac{M}{N}$  is the probability of success each time.

This can ease calculations a lot.

Example: Say 40000 people at a convention, 3000 of them are CEOs, and we meet 15 people in a group (without replacement).

Let  $X$  be the number of people in the group who are CEOs.

Then  $X$  is hypergeometric with  $N = 40000$   $M = 3000$   $n = 15$   
 $N - M = 37000$

$$\text{E.g. } P(X=2) = \frac{\binom{3000}{2} \binom{37000}{13}}{\binom{40000}{15}} \leftarrow \text{hard to do on a handheld calculator}$$
$$= 0.214405 \dots \dots$$

Easier to consider  $Y$  a Binomial ( $n = 15$ ,  $p = \frac{3000}{40000} = \frac{3}{40}$ )

$$P(Y=2) = \binom{15}{2} \left(\frac{3}{40}\right)^2 \left(\frac{37}{40}\right)^{13} = 0.214365 \dots \dots$$

pretty close to  $P(X=2)$  and easier to calculate on a handheld calculator.