# STAT 315: Hypergeometric Random Variables

Luc Rey-Bellet

University of Massachusetts Amherst

*luc@math.umass.edu*

March 4, 2025

# Hypergeometric Random Variable

## Hypergeometric Random Variable

- $N$= size of the population
- Population members come in two kinds: $r$ are of type $I$ and $N - r$ are of type $II$.
- Take a sample of size $n$ in the population. ($n$ must be less than or equal to $N$).
- The hypergeometric random variable $Y$ gives the number of samples which are of type $I$.
- The pdf of $Y$ is

$$p(k) = P(Y = k) = \frac{\binom{r}{k}\binom{N-r}{n-k}}{\binom{N}{n}}$$

The number of possible samples is $\binom{N}{n}$. If $Y = k$ then there are $\binom{r}{k}$ ways for the sample to give $k$ type $I$ and $\binom{N-r}{n-k}$ ways to $n - k$ type $II$.

# Sampling with or without replacement

## Mean of the hypergeometric RV

If $X$ is hypergeomtric with parameters $N$=total population size, $r$=number of type I, $n$=sample size. Then

$$E[X] = \frac{nr}{N}$$

Hypergeometric RV = sampling without replacement: you pick a sample of size $n$ out a population of $N$.

Sample of size $n$ with replacement out of a population of size $N$ with $r$ elements of type $I$ gives a binomial with probability of success $p = \frac{r}{N}$.

$$P(X = k) = \underbrace{\frac{\binom{r}{k}\binom{N-r}{n-k}}{\binom{N}{n}}}_{\text{without replacement}} \quad \text{vs} \quad \underbrace{\binom{n}{k}\left(\frac{r}{N}\right)^k\left(\frac{N-r}{N}\right)^{n-k}}_{\text{with replacement}}$$

# Sampling with or without replacement in picture

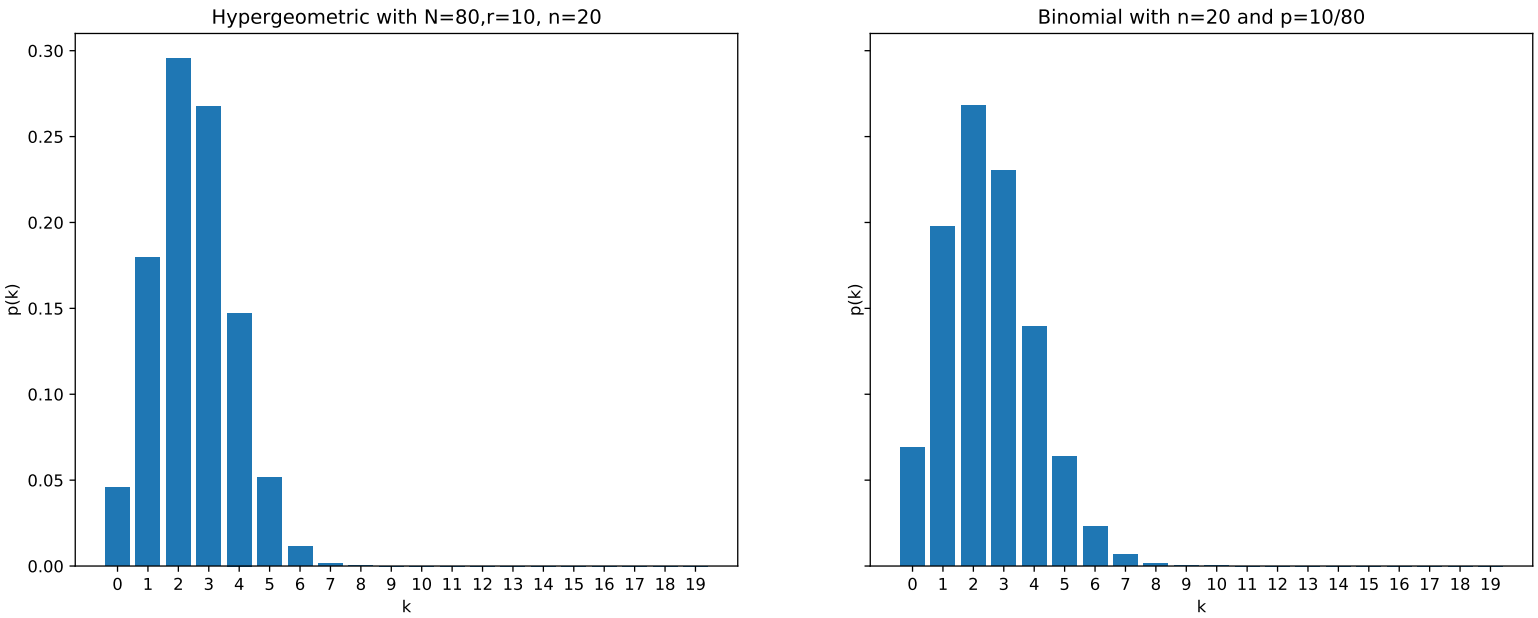


Figure: Left: without replacemnt (hypergeometric). Right: with replacement (binomial). If $N$ is very large one expects that sampling with or without replacement will look quite similar.

# Keno game (often played on video in bars, new draw every few minutes)

Game ticket


- The player select $r$ numbers (for example $r = 10$).
- 20 numbers are drawn from 80 numbers
- Payout if you $k$ of your 10 numbers match.

**How to Win**

Number of Spots

10 Spot Game

| Win Type | Prize* | Probability |
|---|---|---|
| Match 10 | $100,000 | 1:8,911,711.18 |
| Match 9 | $10,000 | 1:163,381.37 |
| Match 8 | $500 | 1:7,384.47 |
| Match 7 | $80 | 1:620.68 |
| Match 6 | $20 | 1:87.11 |
| Match 5 | $2 | 1:19.44 |
| Match 0 | $2 | 1:21.84 |
| **Overall odds are 1 in 9.05** | | |

*See Game Rules for details on prize limitations. All prizes above based on a $1 wager.

# Keno

For a 10 spot game, we have
$N = 80$, $r = 10$ (=the numbers you have chosen)
$n = 20$ sample size, since 20 numbers are sampled out of 80

$$p(k) = P(\text{match of k}) = \frac{\binom{10}{k}\binom{70}{20-k}}{\binom{80}{20}}$$

Expected payout for a bet of \$1.

$$\begin{aligned} E[P] =& 2 \times p(0) + 2 \times p(5) + 20 \times p(6) + 80 \times p(7) + 500 \times p(8) \\ &+ 10,000 \times p(9) + 100,000 \times p(10) = 0.6930 \end{aligned}$$

Ouch.....