# 4 Distribución Hipergeométrica

**Definición 4.1** (Distribución Hipergeométrica)**.** Sea X una v.a discreta, se dice que X sigue una distribución hipergeométrica si su función masa de probabilidad es

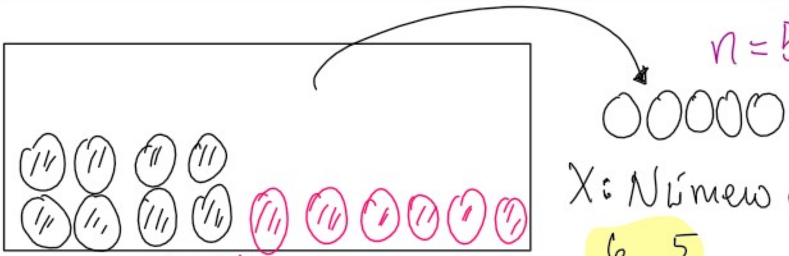
$$\mathbf{P}(X = x) = p(x) = \frac{C_x^r \cdot C_{n-x}^{N-r}}{C_n^N}, \quad \text{máx}\{0, n-(N-r)\} \leqslant x \leqslant \text{mín}\{n, r\} \tag{4.1}$$

$X = 0, 1, 2, \dots, r$

en que:

- $X \equiv$ número de éxitos en la muestra
- $N \equiv$ tamaño de la población
- $n \equiv$ tamaño de la muestra
- $r \equiv$ número de éxitos en la población

**Notación:** $X \sim \text{Hgeo}(N, r, n)$, se lee X sigue una distribución hipergeométrica con parámetros $N, r$ y $n$.



$n = 5$

$\rightsquigarrow X = 0, 1, 2, 3, 4, 5$

X: Número de bolas rojas

$N = 14$

$\frac{6}{14} \cdot \frac{5}{13}$

## Características

- Se tiene que la población N es finita y la selección de la muestras es sin reemplazo
- Los ensayos (o repeticiones) no son independientes y la probabilidad de éxito en cada ensayo no es constante.

El valor esperado y varianza de X son respectivamente

$$\mathbf{E}[X] = n\left(\frac{r}{N}\right) \quad \mathbf{Var}[X] = n\left(\frac{r}{N}\right)\left(1 - \frac{r}{N}\right)\left(\frac{N-n}{N-1}\right)$$

Nótese que $p = r/N$ es la proporción de éxitos en la población, así se puede reescribir

$$\mathbf{E}[X] = np \quad \mathbf{Var}[X] = np\,(1-p)\left(\frac{N-n}{N-1}\right),$$

Si $n$ es menor en relación a N se tiene que

$$\frac{N-n}{N-1} \approx 1 \quad \rightsquigarrow E[X] = np \,; \; Var[X] \approx np(1-p)$$

Notación: $X \sim \text{Hgeo}(N, r, n)$

**Ejemplo 4.1.**
Un subgerente de una empresa de materias primas, debe contratar 10 personas entre 30 candidatos, 22 de los cuales tienen títulos universitarios.

a) Calcule la probabilidad de que 5 de los contratados tengan un título

b) Calcule el número esperado así como la desviación estándar de las personas contratadas que poseen título

**Solución:**
**Éxito:** Se considera como éxito que la persona contratada tenga un título universitario. Así,

- $X \equiv$ número de contratados que tiene título universitario
- $N = 30$, candidatos $\to$ tamaño de la población
- $n = 10$, contratados $\to$ tamaño de la muestra
- $r = 22$, candidatos que poseen título $\to$ número de éxitos en la población

$X \sim Hgeo(30, 22, 10)$

luego la distribución de probabilidad para X es

$$\begin{aligned} \mathbf{P}(X = x) &= \frac{C_x^{22} \cdot C_{10-x}^{30-22}}{C_{10}^{30}} \\ &= \frac{C_x^{22} \cdot C_{10-x}^{8}}{C_{10}^{30}} \end{aligned}$$

en que los valores de X varian

$$\begin{aligned} \max\{0, 10-(30-22)\} \leqslant x &\leqslant \min\{10, 22\} \\ \max\{0, 2\} \leqslant x &\leqslant \min\{10, 22\} \\ 2 \leqslant x &\leqslant 10 \to x = 2, 3, \ldots, 10 \end{aligned}$$

a) la probabilidad de que 5 de los contratados tengan título universitario

$$\mathbf{P}(X = 5) = \frac{C_5^{22} \cdot C_{10-5}^{8}}{C_{10}^{30}} = 0.0491$$

b) el número esperado así como la desviación estándar de las personas contratadas que poseen título

$$\mathbf{E}[X] = n\left(\frac{r}{N}\right) = 10\left(\frac{22}{30}\right) = 7.3333$$

$$\begin{aligned} \sigma = \sqrt{Var[X]} &= \sqrt{n\left(\frac{r}{N}\right)\left(1-\frac{r}{N}\right)\left(\frac{N-n}{N-1}\right)} \\ &= \sqrt{10\left(\frac{22}{30}\right)\left(1-\frac{22}{30}\right)\left(\frac{30-10}{30-1}\right)} \\ &= \end{aligned}$$