

Distribuciones variables aleatorias discretas

Jessica Nathaly Pulzara Mora
jessica.pulzara@udea.edu.co

Departamento de ingeniería de sistemas



**UNIVERSIDAD
DE ANTIOQUIA**

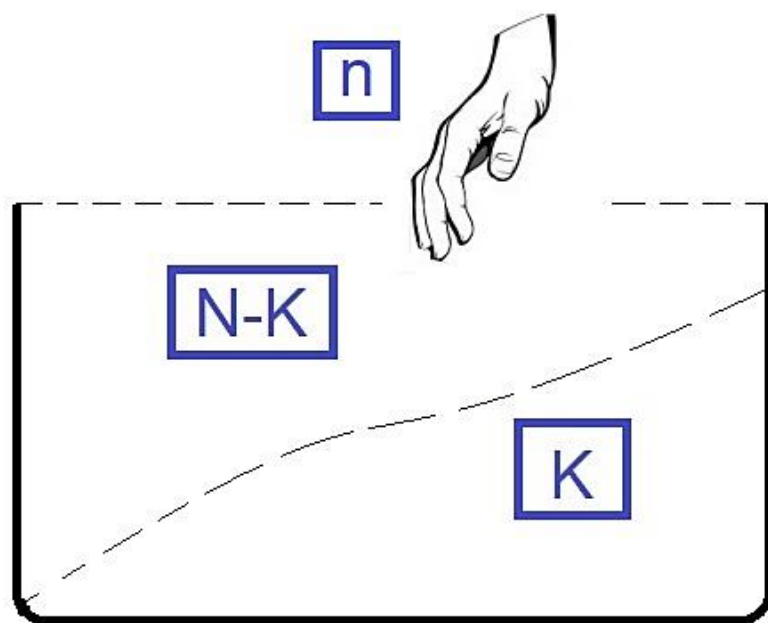
Distribución hipergeométrica

Experimento hipergeométrico

La distribución hipergeométrica modela la probabilidad de los resultados de un experimento así:

- Se tiene una población finita de N elementos.
- Se tienen k éxitos y $N - k$ fracasos.
- Se toman (extrae, eligen) al azar y sin remplazo n de los N elementos.
- Cada elemento tiene sólo dos características posibles: éxito o fracaso (ensayo Bernoulli).
- Por tanto, p (probabilidad de éxito) no es constante y los ensayos no son independientes.

Suponga que una población finita tiene N elementos, cada uno de los cuales tiene una de dos características diferentes (por ejemplo, K elementos tienen la característica de interés y $N - K$ no la tienen). Se toman al azar y sin reemplazo n de estos elementos.



Urna con:

- N : elementos en total
- K : de interés
- $N-K$: no son de interés
- n : elementos seleccionados

Distribución hipergeométrica

- **Variable de interes**

X : número de elementos que tienen la característica de interés en los n seleccionados.

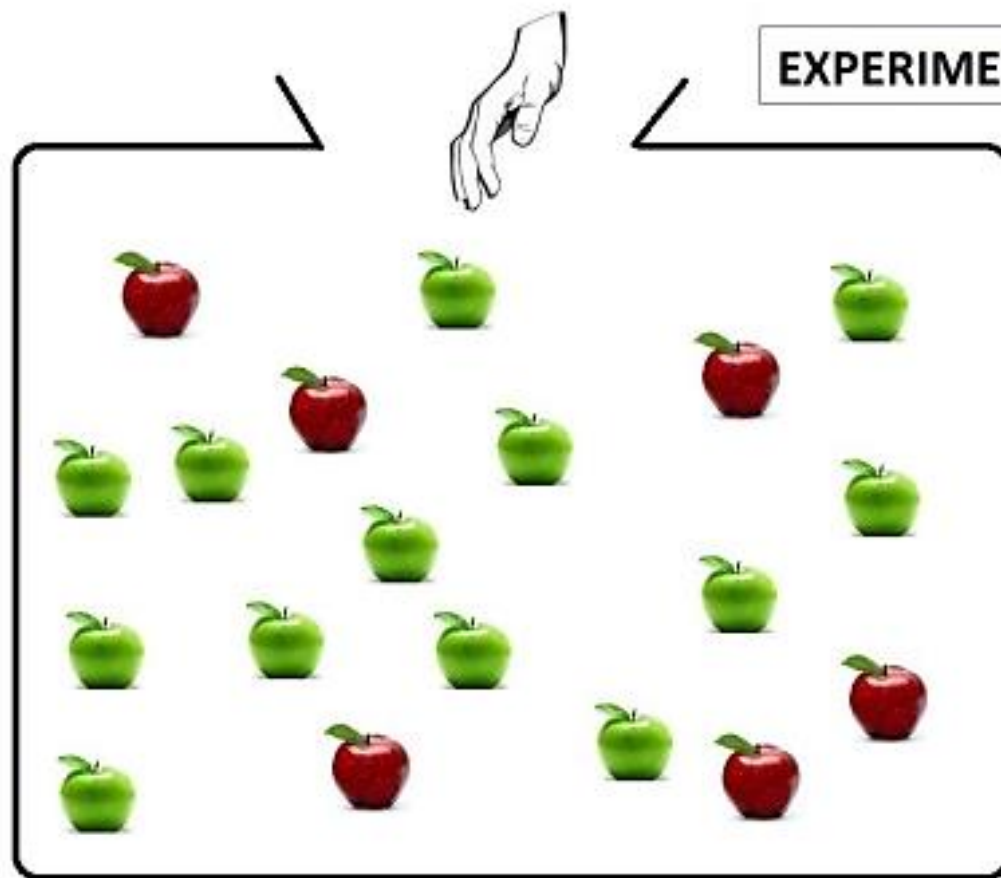
- **Rango**

Los valores que toma esta variable aleatoria son:

$x = 0, 1, 2, \dots, \min(n, K)$

Pero, ¿por qué x va hasta el mínimo entre n y K ?

Pero, ¿por qué x va hasta el mínimo entre n y K ?



EXPERIMENTO: Seleccionar 8 manzanas

**X : # de manzanas rojas
entre las 8
seleccionadas**

$N=19$

$n=8$

$K=6$

$\mathcal{X} = \{0, 1, 2, 3, 4, 5, 6=\min(K, n)\}$

En este experimento se puede observar:

$\binom{N}{n}$: Formas distintas de seleccionar n objetos de N .

$\binom{K}{x}$: Formas distintas de seleccionar x elementos de los K

$\binom{N-K}{n-x}$: Formas distintas de seleccionar $(n-x)$ de los $(N-K)$

Entonces, la f.m.p. de X está dada por:

$$p_X(x) = \begin{cases} \frac{\binom{K}{x} \binom{N-K}{n-x}}{\binom{N}{n}} & ; \quad x = 0, 1, 2, \dots, \min(n, K) \\ 0 & ; \quad \text{e.o.c.} \end{cases}$$

Media y Varianza distribución hipergeométrica

Propiedades: Sea $X \sim \text{Hiper}(N, K, n)$ entonces:

$$\begin{aligned} E[X] &= n \frac{K}{N} \\ \text{Var}[X] &= \frac{(N-n)}{(N-1)} n \frac{K}{N} \left(1 - \frac{K}{N}\right) . \end{aligned}$$

Ejemplo

Un lote con 25 arandelas contiene tres en las que la variabilidad en el espesor alrededor de la circunferencia es inaceptable. Se toma una muestra al azar de tres arandelas.

- ¿Cuál es la probabilidad de que ninguna de las arandelas inaceptables se encuentren en la muestra?
- ¿Cuál es la probabilidad de que al menos una de las arandelas inaceptables se encuentren en la muestra?
- ¿Cuál es la probabilidad de que exactamente una de las arandelas inaceptables se encuentre en la muestra?



Ejemplo

Una persona inescrupulosa va a un banco a consignar 1 millón de pesos en billetes de 50 mil y entre los billetes incluye 10 billetes falsos. Si el cajero del banco solo examina a profundidad 6 billetes:

- ¿Cuál es la f.m.p. para el número de billetes falsos en la muestra?
- ¿Cuál es la probabilidad de que todos los billetes seleccionados sean falsos u originales, es decir, de un solo tipo?
- ¿Cuál es la probabilidad de que la cantidad de billetes falsos seleccionados para ser examinados esté a menos de una desviación estándar de la media?



Comparativo distribuciones binomial e hipergeométrica

Distribution	Population*	Probability of Success (p) [†]	Parameters	
			No. of Trials	No. of Successes
Binomial	Infinite	Constant	Constant (n)	Variable (X)
Hyper-geometric	Finite (K successes; $N - K$ failures)	Changing	Constant (n)	Variable (X)

* If an item selected from a population is replaced before the next trial, the size of the population is considered infinite even if it may be finite.

† If the probability of success p is constant, the trials are considered independent; otherwise, the trials are dependent.

Aproximación binomial de la hipergeométrica

En la distribución binomial, p es constante

Proposición

Sea $X \sim \text{hip}(N, k, n)$:

Si se cumple que

$$\frac{n}{N} < 0.1, \text{ o } \frac{N-n}{N-1} \approx 1$$

entonces la distribución hipergeométrica tiende a la binomial con parámetros n y $p = \frac{k}{N}$:

$$\lim_{N \rightarrow \infty} p(x; N, k, n) = p(x; n, p), \text{ es decir, } \frac{\binom{k}{x} \binom{N-k}{n-x}}{\binom{N}{n}} \approx \binom{n}{x} p^x (1-p)^{n-x}$$

Ejemplo

Considere nuevamente el ejemplo donde una persona inescrupulosa va a un banco a consignar 1 millón de pesos pero ahora en billetes de 5 mil y entre ellos incluye 10 billetes falsos. Si el cajero del banco solo examina a profundidad 6 billetes, ¿cuál es la probabilidad de sacar ningún billete falso o todos los billetes falsos en la muestra?

Sea X : # de billetes falsos en la muestra de 6 billetes.

$$P(X = 0 \vee X = 6) = \frac{\binom{10}{0} \binom{190}{6}}{\binom{200}{6}} + \frac{\binom{10}{6} \binom{190}{0}}{\binom{200}{6}} = 0.7321 .$$

Usando la aproximación se tiene que:

$$P(X = 0 \vee X = 6) \approx \binom{6}{6} \left(\frac{10}{200}\right)^6 \left(\frac{190}{200}\right)^0 + \binom{6}{0} \left(\frac{10}{200}\right)^0 \left(\frac{190}{200}\right)^6 = 0.7351$$

La aproximación es muy buena.

Ejemplo

Se escriben los nombres de 7 mujeres y 3 hombres en pedacitos de papel y se los coloca en una caja. De la caja se escogen al azar 4 papelitos.

- Determine la distribución de probabilidad del número de papelitos seleccionados que contengan nombres de hombres:
 - Si se escogen uno por uno sin reposición
 - si se escogen uno por uno con reposición
- Calcule la probabilidad de que se seleccionen los nombres de por lo menos dos hombres.

Ejercicio

Se tiene 50 representantes de cierto partido político en una convención nacional. 30 de ellos apoyan al candidato A, y 20 al candidato B. Si se seleccionan al azar cinco representantes asistentes a la convención, ¿cuál es la probabilidad de que, entre estos cinco, por lo menos dos apoyen al candidato A?